(12) STANDARD PATENT (11) Application No. **AU 2013201807 C1**
(19) AUSTRALIAN PATENT OFFICE

(54) Title
Amylases, nucleic acids encoding them and methods for making and using them

(51) International Patent Classification(s)
C12N 9/26 (2006.01)
C12N 1/15 (2006.01)
C12N 1/21 (2006.01)
C12N 5/10 (2006.01)
C12N 9/28 (2006.01)
C12N 15/11 (2006.01)
C12N 15/56 (2006.01)
C12N 15/63 (2006.01)
C12Q 1/68 (2006.01)

(21) Application No: **2013201807** (22) Date of Filing: **2013.03.18**

(43) Publication Date: **2013.04.11**
(43) Publication Journal Date: **2013.04.11**
(44) Accepted Journal Date: **2015.09.24**
(44) Amended Journal Date: **2017.11.30**

(62) Divisional of:
2010246342

(71) Applicant(s)
BASF Enzymes LLC

(72) Inventor(s)
Callen, Walter;Richardson, Toby;Frey, Gerhard;Gray, Kevin;Kerovuo, Janne S.;Slupska, Matgorzata;Barton, Nelson;O'Donoghue, Eileen;Miller, Carl

(74) Agent / Attorney
FPA Patent Attorneys Pty Ltd, Level 43 101 Collins Street, Mebourne, VIC, 3000, AU

2013201807 18 Mar 2013

ABSTRACT

AMYLASES, NUCLEIC ACIDS ENCODING THEM AND METHODS FOR MAKING AND USING THEM

In one aspect, the invention is directed to polypeptides having an amylase activity, polynucleotides encoding the polypeptides, and methods for making and using these polynucleotides and polypeptides. In one aspect, the polypeptides of the invention can be used as amylases, for example, alpha amylases, to catalyze the hydrolysis of starch into sugars. In one aspect, the invention provides delayed release compositions comprising an desired ingredient coated by a latex polymer coating.

P/00/011
Regulation 3.2

AUSTRALIA

Patents Act 1990

COMPLETE SPECIFICATION
STANDARD PATENT

Invention Title: **Amylases, nucleic acids encoding them and methods for making and using them**

The following statement is a full description of this invention, including the best method of performing it known to us:

AMYLASES, NUCLEIC ACIDS ENCODING THEM AND METHODS FOR MAKING AND USING THEM

REFERENCE TO SEQUENCE LISTING SUBMITTED ON A COMPACT DISC

This application includes a compact disc (submitted in quadruplicate) containing a sequence listing. The entire content of the sequence listing is herein incorporated by reference. The sequence listing is identified on the compact disc as follows.

File Name	Date of Creation	Size (bytes)
Sequence Listing.txt	March 4, 2004	1,798,144

TECHNICAL FIELD

This invention relates to molecular and cellular biology and biochemistry. In one aspect, the invention is directed to polypeptides having an amylase activity, polynucleotides encoding the polypeptides, and methods for making and using these polynucleotides and polypeptides. In one aspect, the polypeptides of the invention can be used as amylases, for example, alpha amylases or glucoamylases, to catalyze the hydrolysis of starch into sugars. In one aspect, the invention is directed to polypeptides having thermostable amylase activity, such as alpha amylases or glucoamylase activity, e.g., a 1,4-alpha-D-glucan glucohydrolase activity. In one aspect, the polypeptides of the invention can be used as amylases, for example, alpha amylases or glucoamylases, to catalyze the hydrolysis of starch into sugars, such as glucose. The invention is also directed to nucleic acid constructs, vectors, and host cells comprising the nucleic acid sequences of the invention as well as recombinant methods for producing the polypeptides of the invention. The invention is also directed to the use of amylases of the invention in starch conversion processes, including production of high fructose corn syrup (HFCS), ethanol, dextrose, and dextrose syrups.

BACKGROUND

Starch is a complex carbohydrate often found in the human diet. The structure of starch is glucose polymers linked by α-1,4 and α-1,6 glucosidic bonds. Amylase is an enzyme that catalyzes the hydrolysis of starches into sugars. Amylases

hydrolyze internal α-1,4-glucosidic linkages in starch, largely at random, to produce smaller molecular weight malto-dextrins. The breakdown of starch is important in the digestive system and commercially. Amylases are of considerable commercial value, being used in the initial stages (liquefaction) of starch processing; in wet corn milling; in alcohol production; as cleaning agents in detergent matrices; in the textile industry for starch desizing; in baking applications; in the beverage industry; in oilfields in drilling processes; in inking of recycled paper; and in animal feed.

Amylases are produced by a wide variety of microorganisms including *Bacillus* and *Aspergillus*, with most commercial amylases being produced from bacterial sources such as *Bacillus licheniformis*, *Bacillus* amyloliquefaciens, *Bacillus subtilis*, or *Bacillus stearothermophilus*. In recent years, the enzymes in commercial use have been those from *Bacillus licheniformis* because of their heat stability and performance, at least at neutral and mildly alkaline pHs.

Commercially, glucoamylases are used to further hydrolyze cornstarch, which has already been partially hydrolyzed with an alpha-amylase. The glucose produced in this reaction may then be converted to a mixture of glucose and fructose by a glucose isomerase enzyme. This mixture, or one enriched with fructose, is the high fructose corn syrup commercialized throughout the world. In general, starch to fructose processing consists of four steps: liquefaction of granular starch, saccharification of the liquefied starch into dextrose, purification, and isomerization to fructose. The object of a starch liquefaction process is to convert a concentrated suspension of starch polymer granules into a solution of soluble shorter chain length dextrins of low viscosity.

The most widely utilized glucoamylase is produced from the fungus *Aspergillus niger*. One of the problems with the commercial use of this enzyme is its relatively low thermostability. A number of other fungal glucoamylases have been reported, including *Rizopus*, *Thielavia*, *Thermoascus* and *Talaromyces*, and a glucoamylase from the thermophilic fungus *Thermomyces lanuginosus*.

In general, starch to fructose processing consists of four steps: liquefaction of granular starch, saccharification of the liquefied starch into dextrose, purification, and isomerization to fructose. The object of a starch liquefaction process is to convert a concentrated suspension of starch polymer granules into a solution of soluble shorter chain length dextrins of low viscosity. This step is essential for convenient handling with standard equipment and for efficient conversion to glucose or other sugars. To liquefy

granular starch, it is necessary to gelatinize the granules by raising the temperature of the granular starch to over about 72°C. The heating process instantaneously disrupts the insoluble starch granules to produce a water soluble starch solution. The solubilized starch solution is then liquefied by amylase. A starch granule is composed of: 69-74% amylopectin, 26-31% amylose, 11-14% water, 0.2-0.4% protein, 0.5-0.9% lipid, 0.05-0.1% ash, 0.02-0.03% phosphorus, 0.1% pentosan. Approximately 70% of a granule is amorphous and 30% is crystalline.

A common enzymatic liquefaction process involves adjusting the pH of a granular starch slurry to between 6.0 and 6.5, the pH optimum of alpha-amylase derived from *Bacillus licheniformis*, with the addition of calcium hydroxide, sodium hydroxide or sodium carbonate. The addition of calcium hydroxide has the advantage of also providing calcium ions which are known to stabilize the alpha-amylase against inactivation. Upon addition of alpha-amylase, the suspension is pumped through a steam jet to instantaneously raise the temperature to between 80°C to 115°C. The starch is immediately gelatinized and, due to the presence of alpha-amylase, depolymerized through random hydrolysis of a (1-4) glycosidic bonds by alpha-amylase to a fluid mass which is easily pumped.

In a second variation to the liquefaction process, alpha-amylase is added to the starch suspension, the suspension is held at a temperature of 80-100°C to partially hydrolyze the starch granules, and the partially hydrolyzed starch suspension is pumped through a jet at temperatures in excess of about 105°C to thoroughly gelatinize any remaining granular structure. After cooling the gelatinized starch, a second addition of alpha-amylase can be made to further hydrolyze the starch.

A third variation of this process is called the dry milling process. In dry milling, whole grain is ground and combined with water. The germ is optionally removed by flotation separation or equivalent techniques. The resulting mixture, which contains starch, fiber, protein and other components of the grain, is liquefied using alpha-amylase. The general practice in the art is to undertake enzymatic liquefaction at a lower temperature when using the dry milling process. Generally, low temperature liquefaction is believed to be less efficient than high temperature liquefaction in converting starch to soluble dextrins.

Typically, after gelatinization the starch solution is held at an elevated temperature in the presence of alpha-amylase until a DE of 10-20 is achieved, usually a

period of 1-3 hours. Dextrose equivalent (DE) is the industry standard for measuring the concentration of total reducing sugars, calculated as D-glucose on a dry weight basis. Unhydrolyzed granular starch has a DE of virtually zero, whereas the DE of D-glucose is defined as 100.

Corn wet milling is a process which produces corn oil, gluten meal, gluten feed and starch. Alkaline-amylase is used in the liquefaction of starch and glucoamylase is used in saccharification, producing glucose. Corn, a kernel of which consists of a outer seed coat (fiber), starch, a combination of starch and glucose and the inner germ, is subjected to a four step process, which results in the production of starch. The corn is steeped, de-germed, de-fibered, and finally the gluten is separated. In the steeping process, the solubles are taken out. The product remaining after removal of the solubles is de-germed, resulting in production of corn oil and production of an oil cake, which is added to the solubles from the steeping step. The remaining product is de-fibered and the fiber solids are added to the oil cake/solubles mixture. This mixture of fiber solids, oil cake and solubles forms a gluten feed. After de-fibering, the remaining product is subjected to gluten separation. This separation results in a gluten meal and starch. The starch is then subjected to liquefaction and saccharification to produce glucose.

Staling of baked products (such as bread) has been recognized as a problem which becomes more serious as more time lies between the moment of preparation of the bread product and the moment of consumption. The term staling is used to describe changes undesirable to the consumer in the properties of the bread product after leaving the oven, such as an increase of the firmness of the crumb, a decrease of the elasticity of the crumb, and changes in the crust, which becomes tough and leathery. The firmness of the bread crumb increases further during storage up to a level, which is considered as negative. The increase in crumb firmness, which is considered as the most important aspect of staling, is recognized by the consumer a long time before the bread product has otherwise become unsuitable for consumption.

There is a need in the industry for the identification and optimization of amylases, useful for various uses, including commercial cornstarch liquefaction processes. These second generation acid amylases will offer improved manufacturing and/or performance characteristics over the industry standard enzymes from *Bacillus licheniformis*, for example.

There is also a need for the identification and optimization of amylases having utility in automatic dish wash (ADW) products and laundry detergent. In ADW products, the amylase will function at pH 10-11 and at 45-60°C in the presence of calcium chelators and oxidative conditions. For laundry, activity at pH 9-10 and 40°C in the appropriate detergent matrix will be required. Amylases are also useful in textile desizing, brewing processes, starch modification in the paper and pulp industry and other processes described in the art.

Amylases can be used commercially in the initial stages (liquefaction) of starch processing; in wet corn milling; in alcohol production; as cleaning agents in detergent matrices; in the textile industry for starch desizing; in baking applications; in the beverage industry; in oilfields in drilling processes; in inking of recycled paper and in animal feed. Amylases are also useful in textile desizing, brewing processes, starch modification in the paper and pulp industry and other processes.

The publications discussed herein are provided solely for their disclosure prior to the filing date of the present application. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

SUMMARY

The invention provides isolated or recombinant nucleic acids comprising a nucleic acid sequence having at least about 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more, or complete (100%) sequence identity to a nucleic acid of the invention, e.g., an exemplary nucleic acid of the invention, over a region of at least about 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 100, 125, 150, 175, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400, 1450, 1500, 1550 or more, residues. In one aspect, the nucleic acid encodes at least one polypeptide having an amylase activity, and the sequence identities are determined by analysis with a sequence comparison algorithm or by a visual inspection. In another aspect, the invention provides nucleic acids for use as probes, inhibitory molecules (e.g., antisense, iRNAs), transcriptional or translational regulation, and the like. Exemplary

nucleic acids of the invention include isolated or recombinant nucleic acids comprising a nucleic acid sequence as set forth in SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, SEQ ID NO:31, SEQ ID NO:33, SEQ ID NO:35, SEQ ID NO:37, SEQ ID NO:39, SEQ ID NO:41, SEQ ID NO:43, SEQ ID NO:45, SEQ ID NO:47, SEQ ID NO:49, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:55, SEQ ID NO:57, SEQ ID NO:59, SEQ ID NO:61, SEQ ID NO:63, SEQ ID NO:65, SEQ ID NO:67, SEQ ID NO:69, SEQ ID NO:71, SEQ ID NO:73, SEQ ID NO:75, SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81, SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID NO:115, SEQ ID NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID NO:123, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, SEQ ID NO:139, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID NO:147, SEQ ID NO:149, SEQ ID NO:151, SEQ ID NO:153, SEQ ID NO:155, SEQ ID NO:157, SEQ ID NO:159, SEQ ID NO:161, SEQ ID NO:163, SEQ ID NO:165, SEQ ID NO:167, SEQ ID NO:189, SEQ ID NO:191, SEQ ID NO:193, SEQ ID NO:203, SEQ ID NO:205, SEQ ID NO:207, SEQ ID NO:209, SEQ ID NO:211, SEQ ID NO:322, SEQ ID NO:324, SEQ ID NO:326, SEQ ID NO:328, SEQ ID NO:330, SEQ ID NO:332, SEQ ID NO:334, SEQ ID NO:336, SEQ ID NO:338, SEQ ID NO:340, SEQ ID NO:342, SEQ ID NO:344, SEQ ID NO:346, SEQ ID NO:348, SEQ ID NO:350, SEQ ID NO:352, SEQ ID NO:354, SEQ ID NO:356, SEQ ID NO:358, SEQ ID NO:360, SEQ ID NO:362, SEQ ID NO:364, SEQ ID NO:366, SEQ ID NO:368, SEQ ID NO:370, SEQ ID NO:372, SEQ ID NO:374, SEQ ID NO:376, SEQ ID NO:378, SEQ ID NO:380, SEQ ID NO:382, SEQ ID NO:384, SEQ ID NO:386, SEQ ID NO:388, SEQ ID NO:390, SEQ ID NO:392, SEQ ID NO:394, SEQ ID NO:396, SEQ ID NO:398, SEQ ID NO:400, SEQ ID NO:402, SEQ ID NO:404, SEQ ID NO:406, SEQ ID NO:408, SEQ ID NO:410, SEQ ID NO:412, SEQ ID NO:414, SEQ ID NO:416, SEQ ID NO:418, SEQ ID NO:420, SEQ ID NO:422, SEQ ID NO:424, SEQ ID NO:426, SEQ ID NO:428, SEQ ID NO:430, SEQ ID NO:432, SEQ ID NO:434, SEQ ID NO:436, SEQ ID NO:438, SEQ ID NO:440, SEQ ID NO:442, SEQ ID NO:444, SEQ ID NO:446, SEQ ID NO:448, SEQ ID NO:450, SEQ ID NO:452, SEQ ID NO:454, SEQ ID

NO:456, SEQ ID NO:458, SEQ ID NO:460, SEQ ID NO:460, SEQ ID NO:462, SEQ ID NO:465, SEQ ID NO:467, SEQ ID NO:473, SEQ ID NO:475, SEQ ID NO:478, SEQ ID NO:480, SEQ ID NO:484, SEQ ID NO:486, SEQ ID NO:492, SEQ ID NO:494, SEQ ID NO:498, SEQ ID NO:500, SEQ ID NO:509, SEQ ID NO:511, SEQ ID NO:515, SEQ ID NO:517, SEQ ID NO:517, SEQ ID NO:519, SEQ ID NO:522, SEQ ID NO:524, SEQ ID NO:527, SEQ ID NO:529, SEQ ID NO:532, SEQ ID NO:534, SEQ ID NO:539, SEQ ID NO:541, SEQ ID NO:544, SEQ ID NO:546, SEQ ID NO:552, SEQ ID NO:554, SEQ ID NO:558, SEQ ID NO:560, SEQ ID NO:565, SEQ ID NO:567, SEQ ID NO:569, SEQ ID NO:571, SEQ ID NO:573, SEQ ID NO:575, SEQ ID NO:577, SEQ ID NO:579, SEQ ID NO:581, SEQ ID NO:583, SEQ ID NO:585, SEQ ID NO:587, SEQ ID NO:593, SEQ ID NO:603, SEQ ID NO:605, SEQ ID NO:607, SEQ ID NO:609, SEQ ID NO:611, SEQ ID NO:613, SEQ ID NO:615, SEQ ID NO:617, SEQ ID NO:619 or SEQ ID NO:621, and subsequences thereof, e.g., at least about 10, 15, 20, 25, 30, 35, 40, 45, 50, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400, 1450, 1500 or more residues in length, or over the full length of a gene or transcript.

Exemplary nucleic acids of the invention also include isolated or recombinant nucleic acids encoding a polypeptide of the invention, e.g., an exemplary polypeptide having a sequence as set forth in SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:20, SEQ ID NO:22, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32, SEQ ID NO:34, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:44, SEQ ID NO:46, SEQ ID NO:48, SEQ ID NO:50, SEQ ID NO:52, SEQ ID NO:54, SEQ ID NO:56, SEQ ID NO:58, SEQ ID NO:60, SEQ ID NO:62, SEQ ID NO:64, SEQ ID NO:66, SEQ ID NO:68, SEQ ID NO:70, SEQ ID NO:72, SEQ ID NO:74, SEQ ID NO:76, SEQ ID NO:78, SEQ ID NO:80, SEQ ID NO:82, SEQ ID NO:84, SEQ ID NO:86, SEQ ID NO:88, SEQ ID NO:90, SEQ ID NO:92, SEQ ID NO:94, SEQ ID NO:96, SEQ ID NO:98, SEQ ID NO:100, SEQ ID NO:102, SEQ ID NO:104, SEQ ID NO:106, SEQ ID NO:108, SEQ ID NO:110, SEQ ID NO:112, SEQ ID NO:114, SEQ ID NO:116, SEQ ID NO:118, SEQ ID NO:120, SEQ ID NO:122, SEQ ID NO:124, SEQ ID NO:126, SEQ ID NO:128, SEQ ID NO:130, SEQ ID NO:132, SEQ ID NO:134, SEQ ID NO:136, SEQ ID NO:138, SEQ ID NO:140, SEQ ID NO:142, SEQ ID NO:144, SEQ ID NO:146, SEQ ID NO:148, SEQ ID

NO:150, SEQ ID NO:152, SEQ ID NO:154, SEQ ID NO:156, SEQ ID NO:158, SEQ ID NO:160, SEQ ID NO:162, SEQ ID NO:164, SEQ ID NO:166, SEQ ID NO:168, SEQ ID NO:190, SEQ ID NO:192, SEQ ID NO:194, SEQ ID NO:204, SEQ ID NO:206, SEQ ID NO:208, SEQ ID NO:210, SEQ ID NO:212, SEQ ID NO:323, SEQ ID NO:325, SEQ ID NO:327, SEQ ID NO:329, SEQ ID NO:331, SEQ ID NO:333, SEQ ID NO:335, SEQ ID NO:337, SEQ ID NO:339, SEQ ID NO:341, SEQ ID NO:343, SEQ ID NO:345, SEQ ID NO:347, SEQ ID NO:349, SEQ ID NO:351, SEQ ID NO:353, SEQ ID NO:355, SEQ ID NO:357, SEQ ID NO:359, SEQ ID NO:361, SEQ ID NO:363, SEQ ID NO:365, SEQ ID NO:367, SEQ ID NO:369, SEQ ID NO:371, SEQ ID NO:373, SEQ ID NO:375, SEQ ID NO:377, SEQ ID NO:379, SEQ ID NO:381, SEQ ID NO:383, SEQ ID NO:385, SEQ ID NO:387, SEQ ID NO:389, SEQ ID NO:391, SEQ ID NO:393, SEQ ID NO:395, SEQ ID NO:397, SEQ ID NO:399, SEQ ID NO:401, SEQ ID NO:403, SEQ ID NO:405, SEQ ID NO:407, SEQ ID NO:409, SEQ ID NO:411, SEQ ID NO:413, SEQ ID NO:415, SEQ ID NO:417, SEQ ID NO:419, SEQ ID NO:421, SEQ ID NO:423, SEQ ID NO:425, SEQ ID NO:427, SEQ ID NO:429, SEQ ID NO:431, SEQ ID NO:433, SEQ ID NO:435, SEQ ID NO:437, SEQ ID NO:439, SEQ ID NO:441, SEQ ID NO:443, SEQ ID NO:445, SEQ ID NO:447, SEQ ID NO:449, SEQ ID NO:451, SEQ ID NO:453, SEQ ID NO:455, SEQ ID NO:457, SEQ ID NO:459, SEQ ID NO:461, SEQ ID NO:461, SEQ ID NO:463, SEQ ID NO:464, SEQ ID NO:466, SEQ ID NO:468, SEQ ID NO:469, SEQ ID NO:470, SEQ ID NO:471, SEQ ID NO:472, SEQ ID NO:474, SEQ ID NO:476, SEQ ID NO:477, SEQ ID NO:479, SEQ ID NO:481, SEQ ID NO:482, SEQ ID NO:483, SEQ ID NO:485, SEQ ID NO:487, SEQ ID NO:488, SEQ ID NO:489, SEQ ID NO:490, SEQ ID NO:491, SEQ ID NO:493, SEQ ID NO:495, SEQ ID NO:496, SEQ ID NO:497, SEQ ID NO:499, SEQ ID NO:501, SEQ ID NO:502, SEQ ID NO:503, SEQ ID NO:504, SEQ ID NO:505, SEQ ID NO:506, SEQ ID NO:507, SEQ ID NO:508, SEQ ID NO:510, SEQ ID NO:512, SEQ ID NO:513, SEQ ID NO:514, SEQ ID NO:516, SEQ ID NO:518, SEQ ID NO:518, SEQ ID NO:520, SEQ ID NO:521, SEQ ID NO:523, SEQ ID NO:525, SEQ ID NO:526, SEQ ID NO:528, SEQ ID NO:530, SEQ ID NO:531, SEQ ID NO:533, SEQ ID NO:535, SEQ ID NO:536, SEQ ID NO:537, SEQ ID NO:538, SEQ ID NO:540, SEQ ID NO:542, SEQ ID NO:543, SEQ ID NO:545, SEQ ID NO:547, SEQ ID NO:548, SEQ ID NO:549, SEQ ID NO:550, SEQ ID NO:551, SEQ ID NO:553, SEQ ID NO:555, SEQ ID NO:556, SEQ ID NO:557, SEQ ID NO:559, SEQ ID NO:561, SEQ ID NO:562, SEQ ID NO:563, SEQ ID NO:564, SEQ ID NO:566, SEQ ID NO:568, SEQ ID NO:570, SEQ ID NO:572, SEQ ID

NO:574, SEQ ID NO:576, SEQ ID NO:578, SEQ ID NO:580, SEQ ID NO:582, SEQ ID NO:584, SEQ ID NO:586, SEQ ID NO:588, SEQ ID NO:589, SEQ ID NO:590, SEQ ID NO:591, SEQ ID NO:592, SEQ ID NO:594, SEQ ID NO:604, SEQ ID NO:606, SEQ ID NO:608, SEQ ID NO:610, SEQ ID NO:612, SEQ ID NO:614, SEQ ID NO:616, SEQ ID NO:618, SEQ ID NO:620 or SEQ ID NO:622, and subsequences thereof and variants thereof, and polypeptides having at least about 50% (or more, as described below) sequence identity to an exemplary polypeptide of the invention. In one aspect, the polypeptide has an amylase activity, e.g., an alpha amylase or glucoamylase activity (alternative amylase activities described further, below). In one aspect the polypeptide acts as an immunogen or epitope.

In one aspect, the invention also provides amylase-encoding nucleic acids with a common novelty in that they are derived from mixed cultures. The invention provides amylase-encoding nucleic acids isolated from mixed cultures comprising a nucleic acid sequence having at least about 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more, or complete (100%) sequence identity to an exemplary nucleic acid of the invention over a region of at least about 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400, 1450, 1500, 1550 or more, residues, wherein the nucleic acid encodes at least one polypeptide having an amylase activity, and the sequence identities are determined by analysis with a sequence comparison algorithm or by a visual inspection. In one aspect, the invention provides amylase-encoding nucleic acids isolated from mixed cultures comprising a nucleic acid of the invention, e.g., an exemplary nucleic acid of the invention, e.g., a sequence as set forth in SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:9, SEQ ID NO:11, etc., and subsequences thereof, e.g., at least about 10, 15, 20, 25, 30, 35, 40, 45, 50, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400, 1450, 1500 or more residues in length, or over the full length of a gene or transcript; or, a nucleic acid encoding a polypeptide of the invention.

In one aspect, the invention also provides amylase-encoding nucleic acids with a common novelty in that they are derived from environmental sources, e.g., mixed

environmental sources. In one aspect, the invention provides amylase-encoding nucleic acids isolated from environmental sources, e.g., mixed environmental sources, comprising a nucleic acid sequence having at least about 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more, or complete (100%) sequence identity to an exemplary nucleic acid of the invention over a region of at least about 50, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400, 1450, 1500, 1550 or more, residues, wherein the nucleic acid encodes at least one polypeptide having an amylase activity, and the sequence identities are determined by analysis with a sequence comparison algorithm or by a visual inspection. In one aspect, the invention provides amylase-encoding nucleic acids isolated from environmental sources, e.g., mixed environmental sources, comprising a nucleic acid of the invention, e.g., an exemplary nucleic acid sequence of the invention as set forth in SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:9, SEQ ID NO:11, etc., SEQ ID NO:583, SEQ ID NO:585, and subsequences thereof, e.g., at least about 10, 15, 20, 25, 30, 35, 40, 45, 50, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400, 1450, 1500 or more residues in length, or over the full length of a gene or transcript; or, a nucleic acid encoding a polypeptide of the invention.

In one aspect, the invention also provides amylases, and amylase-encoding nucleic acids, with a common novelty in that they are derived from archael sources, including the archael-derived amylases of SEQ ID NO:80 (encoded by SEQ ID NO:79), SEQ ID NO:82 (encoded by SEQ ID NO:81), SEQ ID NO:116 (encoded by SEQ ID NO:115), SEQ ID NO:323 (encoded by SEQ ID NO:322), SEQ NO: 570 (encoded by SEQ ID NO:169).

In one aspect, the sequence comparison algorithm is a BLAST version 2.2.2 algorithm where a filtering setting is set to blastall -p blastp -d "nr pataa" -F F, and all other options are set to default.

Another aspect of the invention is an isolated or recombinant nucleic acid including at least 10 consecutive bases of a nucleic acid sequence of the invention, sequences substantially identical thereto, and the sequences complementary thereto.

In one aspect, the amylase activity comprises a-amylase activity, including the ability to hydrolyze internal alpha-1,4-glucosidic linkages in starch to produce smaller molecular weight malto-dextrins. In one aspect, the a-amylase activity includes hydrolyzing internal alpha-1,4-glucosidic linkages in starch at random. The amylase activity can comprise an a-amylase activity, a ß-amylase activity, a glucoamylase activity, a 1,4-a-D-glucan glucohydrolase activity, an exoamylase activity, a glucan a-maltotetrahydrolase activity, a maltase activity, an isomaltase activity, a glucan 1, 4, a-glucosidase activity, an a-glucosidase activity, a sucrase activity or an agarase activity (e.g., a ß-agarase activity).

The amylase activity can comprise hydrolyzing glucosidic bonds. In one aspect, the glucosidic bonds comprise an a-1,4-glucosidic bond. In another aspect, the glucosidic bonds comprise an a-1,6-glucosidic bond. In one aspect, the amylase activity comprises hydrolyzing glucosidic bonds in starch, e.g., liquefied starch. The amylase activity can further comprise hydrolyzing glucosidic bonds into maltodextrins. In one aspect, the amylase activity comprises cleaving a maltose or a D-glucose unit from non-reducing end of the starch.

In one aspect, the isolated or recombinant nucleic acid encodes a polypeptide having an amylase activity which is thermostable. The polypeptide can retain an amylase activity under conditions comprising a temperature range of anywhere between about 0°C to about 37°C, or, between about 37°C to about 95°C or more, e.g., 98°C, 100°C or more; between about 55°C to about 85°C, between about 70°C to about 95°C, or, between about 90°C to about 95°C. For example, the exemplary polypeptide having a sequence as set forth in SEQ ID NO:437 is thermostable, retaining 50% activity after 25 minutes at 100°C in the absence of added calcium.

In another aspect, the isolated or recombinant nucleic acid encodes a polypeptide having an amylase activity which is thermotolerant. The polypeptide can retain an amylase activity after exposure to a temperature in the range from greater than 37°C to about 95°C or anywhere in the range from greater than 55°C to about 85°C. In one aspect, the polypeptide retains an amylase activity after exposure to a temperature in the range from greater than 90°C to about 95°C at pH 4.5.

The invention provides isolated or recombinant nucleic acids comprising a sequence that hybridizes under stringent conditions to a nucleic acid of the invention, e.g., an exemplary nucleic acid of the invention, a nucleic acid comprising a sequence as set

forth in SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, SEQ ID NO:31, SEQ ID NO:33, SEQ ID NO:35, SEQ ID NO:37, SEQ ID NO:39, SEQ ID NO:41, SEQ ID NO:43, SEQ ID NO:45, SEQ ID NO:47, SEQ ID NO:49, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:55, SEQ ID NO:57, SEQ ID NO:59, SEQ ID NO:61, SEQ ID NO:63, SEQ ID NO:65, SEQ ID NO:67, SEQ ID NO:69, SEQ ID NO:71, SEQ ID NO:73, SEQ ID NO:75, SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81, SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID NO:115, SEQ ID NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID NO:123, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, SEQ ID NO:139, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID NO:147, SEQ ID NO:149, SEQ ID NO:151, SEQ ID NO:153, SEQ ID NO:155, SEQ ID NO:157, SEQ ID NO:159, SEQ ID NO:161, SEQ ID NO:163, SEQ ID NO:165, SEQ ID NO:167, SEQ ID NO:189, SEQ ID NO:191, SEQ ID NO:193, SEQ ID NO:203, SEQ ID NO:205, SEQ ID NO:207, SEQ ID NO:209, SEQ ID NO:211, SEQ ID NO:322, SEQ ID NO:324, SEQ ID NO:326, SEQ ID NO:328, SEQ ID NO:330, SEQ ID NO:332, SEQ ID NO:334, SEQ ID NO:336, SEQ ID NO:338, SEQ ID NO:340, SEQ ID NO:342, SEQ ID NO:344, SEQ ID NO:346, SEQ ID NO:348, SEQ ID NO:350, SEQ ID NO:352, SEQ ID NO:354, SEQ ID NO:356, SEQ ID NO:358, SEQ ID NO:360, SEQ ID NO:362, SEQ ID NO:364, SEQ ID NO:366, SEQ ID NO:368, SEQ ID NO:370, SEQ ID NO:372, SEQ ID NO:374, SEQ ID NO:376, SEQ ID NO:378, SEQ ID NO:380, SEQ ID NO:382, SEQ ID NO:384, SEQ ID NO:386, SEQ ID NO:388, SEQ ID NO:390, SEQ ID NO:392, SEQ ID NO:394, SEQ ID NO:396, SEQ ID NO:398, SEQ ID NO:400, SEQ ID NO:402, SEQ ID NO:404, SEQ ID NO:406, SEQ ID NO:408, SEQ ID NO:410, SEQ ID NO:412, SEQ ID NO:414, SEQ ID NO:416, SEQ ID NO:418, SEQ ID NO:420, SEQ ID NO:422, SEQ ID NO:424, SEQ ID NO:426, SEQ ID NO:428, SEQ ID NO:430, SEQ ID NO:432, SEQ ID NO:434, SEQ ID NO:436, SEQ ID NO:438, SEQ ID NO:440, SEQ ID NO:442, SEQ ID NO:444, SEQ ID NO:446, SEQ ID NO:448, SEQ ID NO:450, SEQ ID NO:452, SEQ ID NO:454, SEQ ID NO:456, SEQ ID NO:458, SEQ ID NO:460, SEQ ID NO:460, SEQ ID NO:462, SEQ ID NO:465, SEQ ID NO:467, SEQ ID

NO:473, SEQ ID NO:475, SEQ ID NO:478, SEQ ID NO:480, SEQ ID NO:484, SEQ ID NO:486, SEQ ID NO:492, SEQ ID NO:494, SEQ ID NO:498, SEQ ID NO:500, SEQ ID NO:509, SEQ ID NO:511, SEQ ID NO:515, SEQ ID NO:517, SEQ ID NO:517, SEQ ID NO:519, SEQ ID NO:522, SEQ ID NO:524, SEQ ID NO:527, SEQ ID NO:529, SEQ ID NO:532, SEQ ID NO:534, SEQ ID NO:539, SEQ ID NO:541, SEQ ID NO:544, SEQ ID NO:546, SEQ ID NO:552, SEQ ID NO:554, SEQ ID NO:558, SEQ ID NO:560, SEQ ID NO:565, SEQ ID NO:567, SEQ ID NO:569, SEQ ID NO:571, SEQ ID NO:573, SEQ ID NO:575, SEQ ID NO:577, SEQ ID NO:579, SEQ ID NO:581, SEQ ID NO:583, SEQ ID NO:585, SEQ ID NO:587, SEQ ID NO:593, SEQ ID NO:603, SEQ ID NO:605, SEQ ID NO:607, SEQ ID NO:609, SEQ ID NO:611, SEQ ID NO:613, SEQ ID NO:615, SEQ ID NO:617, SEQ ID NO:619 or SEQ ID NO:621, or fragments or subsequences thereof. In one aspect, the nucleic acid encodes a polypeptide having an amylase activity. The nucleic acid can be at least about 50, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400, 1450, 1500 or more residues in length or the full length of the gene or transcript. In one aspect, the stringent conditions include a wash step comprising a wash in 0.2X SSC at a temperature of about 65°C for about 15 minutes.

The invention provides a nucleic acid probe for identifying a nucleic acid encoding a polypeptide having an amylase activity, wherein the probe comprises at least about 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000 or more, consecutive bases of a sequence comprising a sequence of the invention, or fragments or subsequences thereof, wherein the probe identifies the nucleic acid by binding or hybridization. The probe can comprise an oligonucleotide comprising at least about 10 to 50, about 20 to 60, about 30 to 70, about 40 to 80, or about 60 to 100 consecutive bases of a sequence comprising a sequence of the invention, or fragments or subsequences thereof.

The invention provides a nucleic acid probe for identifying a nucleic acid encoding a polypeptide having an amylase activity, wherein the probe comprises a nucleic acid comprising a sequence at least about 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 90, 100, 125, 150, 175, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000 or more residues having at least about 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%,

82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more, or complete (100%) sequence identity to a nucleic acid of the invention, wherein the sequence identities are determined by analysis with a sequence comparison algorithm or by visual inspection.

The probe can comprise an oligonucleotide comprising at least about 10 to 50, about 20 to 60, about 30 to 70, about 40 to 80, or about 60 to 100 consecutive bases of a nucleic acid sequence of the invention, or a subsequence thereof.

The invention provides an amplification primer sequence pair for amplifying a nucleic acid encoding a polypeptide having an amylase activity, wherein the primer pair is capable of amplifying a nucleic acid comprising a sequence of the invention, or fragments or subsequences thereof. One or each member of the amplification primer sequence pair can comprise an oligonucleotide comprising at least about 10 to 50 consecutive bases of the sequence.

The invention provides methods of amplifying a nucleic acid encoding a polypeptide having an amylase activity comprising amplification of a template nucleic acid with an amplification primer sequence pair capable of amplifying a nucleic acid sequence of the invention, or fragments or subsequences thereof.

The invention provides expression cassettes comprising a nucleic acid of the invention or a subsequence thereof. In one aspect, the expression cassette can comprise the nucleic acid that is operably linked to a promoter. The promoter can be a viral, bacterial, mammalian or plant promoter. In one aspect, the plant promoter can be a potato, rice, corn, wheat, tobacco or barley promoter. The promoter can be a constitutive promoter. The constitutive promoter can comprise CaMV35S. In another aspect, the promoter can be an inducible promoter. In one aspect, the promoter can be a tissue-specific promoter or an environmentally regulated or a developmentally regulated promoter. Thus, the promoter can be, e.g., a seed-specific, a leaf-specific, a root-specific, a stem-specific or an abscission-induced promoter. In one aspect, the expression cassette can further comprise a plant or plant virus expression vector.

The invention provides cloning vehicles comprising an expression cassette (e.g., a vector) of the invention or a nucleic acid of the invention. The cloning vehicle can be a viral vector, a plasmid, a phage, a phagemid, a cosmid, a fosmid, a bacteriophage or an artificial chromosome. The viral vector can comprise an adenovirus vector, a retroviral vector or an adeno-associated viral vector. The cloning vehicle can comprise a

bacterial artificial chromosome (BAC), a plasmid, a bacteriophage P1-derived vector (PAC), a yeast artificial chromosome (YAC), or a mammalian artificial chromosome (MAC).

The invention provides transformed cell comprising a nucleic acid of the invention or an expression cassette (e.g., a vector) of the invention, or a cloning vehicle of the invention. In one aspect, the transformed cell can be a bacterial cell, a mammalian cell, a fungal cell, a yeast cell, an insect cell or a plant cell. In one aspect, the plant cell can be a potato, wheat, rice, corn, tobacco or barley cell.

The invention provides transgenic non-human animals comprising a nucleic acid of the invention or an expression cassette (e.g., a vector) of the invention. In one aspect, the animal is a mouse.

The invention provides transgenic plants comprising a nucleic acid of the invention or an expression cassette (e.g., a vector) of the invention. The transgenic plant can be a corn plant, a potato plant, a tomato plant, a wheat plant, an oilseed plant, a rapeseed plant, a soybean plant, a rice plant, a barley plant or a tobacco plant.

The invention provides transgenic seeds comprising a nucleic acid of the invention or an expression cassette (e.g., a vector) of the invention. The transgenic seed can be a corn seed, a wheat kernel, an oilseed, a rapeseed, a soybean seed, a palm kernel, a sunflower seed, a sesame seed, a peanut or a tobacco plant seed.

The invention provides an antisense oligonucleotide comprising a nucleic acid sequence complementary to or capable of hybridizing under stringent conditions to a nucleic acid of the invention. The invention provides methods of inhibiting the translation of an amylase message in a cell comprising administering to the cell or expressing in the cell an antisense oligonucleotide comprising a nucleic acid sequence complementary to or capable of hybridizing under stringent conditions to a nucleic acid of the invention.

The invention provides an isolated or recombinant polypeptide comprising an amino acid sequence having at least about 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more, or complete (100%) sequence identity to an exemplary polypeptide or peptide of the invention over a region of at least about 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70,

75, 80, 90, 100, 125, 150, 175, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400, 1450, 1500, 1550 or more residues, or over the full length of the polypeptide, and the sequence identities are determined by analysis with a sequence comparison algorithm or by a visual inspection. Exemplary polypeptide or peptide sequences of the invention include SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:20, SEQ ID NO:22, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32, SEQ ID NO:34, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:44, SEQ ID NO:46, SEQ ID NO:48, SEQ ID NO:50, SEQ ID NO:52, SEQ ID NO:54, SEQ ID NO:56, SEQ ID NO:58, SEQ ID NO:60, SEQ ID NO:62, SEQ ID NO:64, SEQ ID NO:66, SEQ ID NO:68, SEQ ID NO:70, SEQ ID NO:72, SEQ ID NO:74, SEQ ID NO:76, SEQ ID NO:78, SEQ ID NO:80, SEQ ID NO:82, SEQ ID NO:84, SEQ ID NO:86, SEQ ID NO:88, SEQ ID NO:90, SEQ ID NO:92, SEQ ID NO:94, SEQ ID NO:96, SEQ ID NO:98, SEQ ID NO:100, SEQ ID NO:102, SEQ ID NO:104, SEQ ID NO:106, SEQ ID NO:108, SEQ ID NO:110, SEQ ID NO:112, SEQ ID NO:114, SEQ ID NO:116, SEQ ID NO:118, SEQ ID NO:120, SEQ ID NO:122, SEQ ID NO:124, SEQ ID NO:126, SEQ ID NO:128, SEQ ID NO:130, SEQ ID NO:132, SEQ ID NO:134, SEQ ID NO:136, SEQ ID NO:138, SEQ ID NO:140, SEQ ID NO:142, SEQ ID NO:144, SEQ ID NO:146, SEQ ID NO:148, SEQ ID NO:150, SEQ ID NO:152, SEQ ID NO:154, SEQ ID NO:156, SEQ ID NO:158, SEQ ID NO:160, SEQ ID NO:162, SEQ ID NO:164, SEQ ID NO:166, SEQ ID NO:168, SEQ ID NO:190, SEQ ID NO:192, SEQ ID NO:194, SEQ ID NO:204, SEQ ID NO:206, SEQ ID NO:208, SEQ ID NO:210, SEQ ID NO:212, SEQ ID NO:323, SEQ ID NO:325, SEQ ID NO:327, SEQ ID NO:329, SEQ ID NO:331, SEQ ID NO:333, SEQ ID NO:335, SEQ ID NO:337, SEQ ID NO:339, SEQ ID NO:341, SEQ ID NO:343, SEQ ID NO:345, SEQ ID NO:347, SEQ ID NO:349, SEQ ID NO:351, SEQ ID NO:353, SEQ ID NO:355, SEQ ID NO:357, SEQ ID NO:359, SEQ ID NO:361, SEQ ID NO:363, SEQ ID NO:365, SEQ ID NO:367, SEQ ID NO:369, SEQ ID NO:371, SEQ ID NO:373, SEQ ID NO:375, SEQ ID NO:377, SEQ ID NO:379, SEQ ID NO:381, SEQ ID NO:383, SEQ ID NO:385, SEQ ID NO:387, SEQ ID NO:389, SEQ ID NO:391, SEQ ID NO:393, SEQ ID NO:395, SEQ ID NO:397, SEQ ID NO:399, SEQ ID NO:401, SEQ ID NO:403, SEQ ID NO:405, SEQ ID NO:407, SEQ ID NO:409, SEQ ID NO:411, SEQ ID NO:413, SEQ ID NO:415, SEQ ID NO:417, SEQ ID NO:419, SEQ ID NO:421, SEQ ID

NO:423, SEQ ID NO:425, SEQ ID NO:427, SEQ ID NO:429, SEQ ID NO:431, SEQ ID NO:433, SEQ ID NO:435, SEQ ID NO:437, SEQ ID NO:439, SEQ ID NO:441, SEQ ID NO:443, SEQ ID NO:445, SEQ ID NO:447, SEQ ID NO:449, SEQ ID NO:451, SEQ ID NO:453, SEQ ID NO:455, SEQ ID NO:457, SEQ ID NO:459, SEQ ID NO:461, SEQ ID NO:461, SEQ ID NO:463, SEQ ID NO:464, SEQ ID NO:466, SEQ ID NO:468, SEQ ID NO:469, SEQ ID NO:470, SEQ ID NO:471, SEQ ID NO:472, SEQ ID NO:474, SEQ ID NO:476, SEQ ID NO:477, SEQ ID NO:479, SEQ ID NO:481, SEQ ID NO:482, SEQ ID NO:483, SEQ ID NO:485, SEQ ID NO:487, SEQ ID NO:488, SEQ ID NO:489, SEQ ID NO:490, SEQ ID NO:491, SEQ ID NO:493, SEQ ID NO:495, SEQ ID NO:496, SEQ ID NO:497, SEQ ID NO:499, SEQ ID NO:501, SEQ ID NO:502, SEQ ID NO:503, SEQ ID NO:504, SEQ ID NO:505, SEQ ID NO:506, SEQ ID NO:507, SEQ ID NO:508, SEQ ID NO:510, SEQ ID NO:512, SEQ ID NO:513, SEQ ID NO:514, SEQ ID NO:516, SEQ ID NO:518, SEQ ID NO:518, SEQ ID NO:520, SEQ ID NO:521, SEQ ID NO:523, SEQ ID NO:525, SEQ ID NO:526, SEQ ID NO:528, SEQ ID NO:530, SEQ ID NO:531, SEQ ID NO:533, SEQ ID NO:535, SEQ ID NO:536, SEQ ID NO:537, SEQ ID NO:538, SEQ ID NO:540, SEQ ID NO:542, SEQ ID NO:543, SEQ ID NO:545, SEQ ID NO:547, SEQ ID NO:548, SEQ ID NO:549, SEQ ID NO:550, SEQ ID NO:551, SEQ ID NO:553, SEQ ID NO:555, SEQ ID NO:556, SEQ ID NO:557, SEQ ID NO:559, SEQ ID NO:561, SEQ ID NO:562, SEQ ID NO:563, SEQ ID NO:564, SEQ ID NO:566, SEQ ID NO:568, SEQ ID NO:570, SEQ ID NO:572, SEQ ID NO:574, SEQ ID NO:576, SEQ ID NO:578, SEQ ID NO:580, SEQ ID NO:582, SEQ ID NO:584, SEQ ID NO:586, SEQ ID NO:588, SEQ ID NO:589, SEQ ID NO:590, SEQ ID NO:591, SEQ ID NO:592, SEQ ID NO:594, SEQ ID NO:604, SEQ ID NO:606, SEQ ID NO:608, SEQ ID NO:610, SEQ ID NO:612, SEQ ID NO:614, SEQ ID NO:616, SEQ ID NO:618, SEQ ID NO:620 or SEQ ID NO:622, and subsequences thereof and variants thereof, e.g., at least about 10, 15, 20, 25, 30, 35, 40, 45, 50, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400, 1450, 1500 or more residues in length, or over the full length of an enzyme. Exemplary polypeptide or peptide sequences of the invention include sequence encoded by a nucleic acid of the invention. Exemplary polypeptide or peptide sequences of the invention include polypeptides or peptides specifically bound by an antibody of the invention. In one aspect, a polypeptide of the invention has at least one amylase activity, e.g., an alpha amylase activity.

Another aspect of the invention is an isolated or recombinant polypeptide or peptide including at least 10 consecutive bases of a polypeptide or peptide sequence of the invention, sequences substantially identical thereto, and the sequences complementary thereto.

In one aspect, the amylase activity of a polypeptide or peptide of the invention comprises an a-amylase activity, including the ability to hydrolyze internal alpha-1,4-glucosidic linkages in starch to produce smaller molecular weight malto-dextrins. In one aspect, the a-amylase activity includes hydrolyzing internal alpha-1,4-glucosidic linkages in starch at random. The amylase activity can comprise a glucoamylase activity, a 1,4-a-D-glucan glucohydrolase activity, an a-amylase activity, an exoamylase activity, or a ß-amylase activity. The amylase activity can comprise hydrolyzing glucosidic bonds. In one aspect, the glucosidic bonds comprise an a-1,4-glucosidic bond. In another aspect, the glucosidic bonds comprise an a-1,6-glucosidic bond. In one aspect, the amylase activity comprises hydrolyzing glucosidic bonds in starch, e.g., liquefied starch. The amylase activity can further comprise hydrolyzing glucosidic bonds into maltodextrins. In one aspect, the amylase activity comprises cleaving a maltose or a D-glucose unit from non-reducing end of the starch.

In one aspect, the amylase activity of the invention comprises a glucoamylase activity, which can comprise catalysis of the hydrolysis of glucosidic bonds. The glucoamylase activity of the invention can comprise catalyzing the step-wise hydrolytic release of D-glucose from the non-reducing ends of starch or other related dextrins. The glucoamylase activity can comprise a 1,4-a-D-glucan glucohydralase activity. The glucoamylase activity can comprise catalysis of the hydrolysis of malto-dextrins resulting in the generation of free glucose. The glucoamylase activity can comprise an exoamylase activity. The glucoamylase activity can comprise an a-amylase or a ß-amylase activity. The hydrolyzed glucosidic bonds can comprise a-1,4-glucosidic bonds or a-1,6-glucosidic bonds. The glucoamylase activity can comprise hydrolyzing glucosidic bonds in a starch. The glucoamylase activity can further comprise hydrolyzing glucosidic bonds in the starch to produce maltodextrines. The glucoamylase activity can comprise cleaving a maltose or a D-glucose unit from non-reducing end of the starch.

In one aspect, the amylase activity can be thermostable. The polypeptide can retain an amylase activity under conditions comprising a temperature range of between about 37°C to about 95°C, between about 55°C to about 85°C, between about

70°C to about 95°C, or between about 90°C to about 95°C. In another aspect, the amylase activity can be thermotolerant. The polypeptide can retain an amylase activity after exposure to a temperature in the range from greater than 37°C to about 95°C, or in the range from greater than 55°C to about 85°C. In one aspect, the polypeptide can retain an amylase activity after exposure to a temperature in the range from greater than 90°C to about 95°C at pH 4.5.

In one aspect, the isolated or recombinant polypeptide can comprise the polypeptide of the invention that lacks a signal sequence. In one aspect, the isolated or recombinant polypeptide can comprise the polypeptide of the invention comprising a heterologous signal sequence, such as a heterologous amylase or non-amylase signal sequence.

In one aspect, the invention provides a signal sequence comprising a peptide as set forth in Table 3. In one aspect, the invention provides a signal sequence consisting of a peptide as set forth in Table 3. In one aspect, the invention provides chimeric proteins comprising a first domain comprising a signal sequence of the invention and at least a second domain. The protein can be a fusion protein. The second domain can comprise an enzyme. The enzyme can be an amylase (e.g., an amylase of the invention, or, another amylase).

In one aspect, the amylase activity comprises a specific activity at about 37°C in the range from about 10 to 10,000, or, 100 to about 1000 units per milligram of protein. In another aspect, the amylase activity comprises a specific activity from about 500 to about 750 units per milligram of protein. Alternatively, the amylase activity comprises a specific activity at 37°C in the range from about 500 to about 1200 units per milligram of protein. In one aspect, the amylase activity comprises a specific activity at 37°C in the range from about 750 to about 1000 units per milligram of protein. In another aspect, the thermotolerance comprises retention of at least half of the specific activity of the amylase at 37°C after being heated to the elevated temperature. Alternatively, the thermotolerance can comprise retention of specific activity at 37°C in the range from about 500 to about 1200 units per milligram of protein after being heated to the elevated temperature.

The invention provides isolated or recombinant polypeptides of the invention, wherein the polypeptide comprises at least one glycosylation site. In one aspect, glycosylation can be an N-linked glycosylation. In one aspect, the polypeptide

can be glycosylated after being expressed in a *P. pastoris* or a *S. pombe*. The invention also provides methods for adding glycosylation to a polypeptide, either post-translationally or chemically, to change the property of the polypeptides, e.g., its thermal stability, solubility, tendency to aggregate, and the like.

In one aspect, the polypeptide can retain an amylase activity under conditions comprising about pH 6.5, pH 6, pH 5.5, pH 5, pH 4.5 or pH 4. In another aspect, the polypeptide can retain an amylase activity under conditions comprising about pH 7, pH 7.5 pH 8.0, pH 8.5, pH 9, pH 9.5, pH 10, pH 10.5 or pH 11.

The invention provides protein preparations comprising a polypeptide of the invention, wherein the protein preparation comprises a liquid, a solid or a gel.

The invention provides heterodimers comprising a polypeptide of the invention and a second domain. In one aspect, the second domain can be a polypeptide and the heterodimer can be a fusion protein. In one aspect, the second domain can be an epitope or a tag. In one aspect, the invention provides homodimers comprising a polypeptide of the invention.

The invention provides immobilized polypeptides having an amylase activity, wherein the polypeptide comprises a polypeptide of the invention, a polypeptide encoded by a nucleic acid of the invention, or a polypeptide comprising a polypeptide of the invention and a second domain. In one aspect, the polypeptide can be immobilized on a cell, a metal, a resin, a polymer, a ceramic, a glass, a microelectrode, a graphitic particle, a bead, a gel, a plate, an array or a capillary tube.

The invention provides arrays comprising an immobilized nucleic acid of the invention. The invention provides arrays comprising an antibody of the invention.

The invention provides isolated or recombinant antibodies that specifically bind to a polypeptide of the invention or to a polypeptide encoded by a nucleic acid of the invention. The antibody can be a monoclonal or a polyclonal antibody. The invention provides hybridomas comprising an antibody of the invention, e.g., an antibody that specifically binds to a polypeptide of the invention or to a polypeptide encoded by a nucleic acid of the invention.

The invention provides food supplements for an animal comprising a polypeptide of the invention, e.g., a polypeptide encoded by the nucleic acid of the invention. In one aspect, the polypeptide in the food supplement can be glycosylated. The invention provides edible enzyme delivery matrices comprising a polypeptide of the

invention, e.g., a polypeptide encoded by the nucleic acid of the invention. In one aspect, the delivery matrix comprises a pellet. In one aspect, the polypeptide can be glycosylated. In one aspect, the amylase activity is thermotolerant. In another aspect, the amylase activity is thermostable.

The invention provides method of isolating or identifying a polypeptide having an amylase activity comprising the steps of: (a) providing an antibody of the invention; (b) providing a sample comprising polypeptides; and (c) contacting the sample of step (b) with the antibody of step (a) under conditions wherein the antibody can specifically bind to the polypeptide, thereby isolating or identifying a polypeptide having an amylase activity.

The invention provides methods of making an anti-amylase antibody comprising administering to a non-human animal a nucleic acid of the invention or a polypeptide of the invention or subsequences thereof in an amount sufficient to generate a humoral immune response, thereby making an anti-amylase antibody. The invention provides methods of making an anti-amylase immune comprising administering to a non-human animal a nucleic acid of the invention or a polypeptide of the invention or subsequences thereof in an amount sufficient to generate an immune response.

The invention provides methods of producing a recombinant polypeptide comprising the steps of: (a) providing a nucleic acid of the invention operably linked to a promoter; and (b) expressing the nucleic acid of step (a) under conditions that allow expression of the polypeptide, thereby producing a recombinant polypeptide. In one aspect, the method can further comprise transforming a host cell with the nucleic acid of step (a) followed by expressing the nucleic acid of step (a), thereby producing a recombinant polypeptide in a transformed cell.

The invention provides methods for identifying a polypeptide having an amylase activity comprising the following steps: (a) providing a polypeptide of the invention; or a polypeptide encoded by a nucleic acid of the invention; (b) providing an amylase substrate; and (c) contacting the polypeptide or a fragment or variant thereof of step (a) with the substrate of step (b) and detecting a decrease in the amount of substrate or an increase in the amount of a reaction product, wherein a decrease in the amount of the substrate or an increase in the amount of the reaction product detects a polypeptide having an amylase activity. In one aspect, the substrate can be a starch, e.g., a liquefied starch.

The invention provides methods for identifying an amylase substrate comprising the following steps: (a) providing a polypeptide of the invention; or a polypeptide encoded by a nucleic acid of the invention; (b) providing a test substrate; and (c) contacting the polypeptide of step (a) with the test substrate of step (b) and detecting a decrease in the amount of substrate or an increase in the amount of reaction product, wherein a decrease in the amount of the substrate or an increase in the amount of a reaction product identifies the test substrate as an amylase substrate.

The invention provides methods of determining whether a test compound specifically binds to a polypeptide comprising the following steps: (a) expressing a nucleic acid or a vector comprising the nucleic acid under conditions permissive for translation of the nucleic acid to a polypeptide, wherein the nucleic acid comprises a nucleic acid of the invention, or, providing a polypeptide of the invention; (b) providing a test compound; (c) contacting the polypeptide with the test compound; and (d) determining whether the test compound of step (b) specifically binds to the polypeptide.

The invention provides methods for identifying a modulator of an amylase activity comprising the following steps: (a) providing a polypeptide of the invention or a polypeptide encoded by a nucleic acid of the invention; (b) providing a test compound; (c) contacting the polypeptide of step (a) with the test compound of step (b) and measuring an activity of the amylase, wherein a change in the amylase activity measured in the presence of the test compound compared to the activity in the absence of the test compound provides a determination that the test compound modulates the amylase activity. In one aspect, the amylase activity can be measured by providing an amylase substrate and detecting a decrease in the amount of the substrate or an increase in the amount of a reaction product, or, an increase in the amount of the substrate or a decrease in the amount of a reaction product. A decrease in the amount of the substrate or an increase in the amount of the reaction product with the test compound as compared to the amount of substrate or reaction product without the test compound identifies the test compound as an activator of amylase activity. An increase in the amount of the substrate or a decrease in the amount of the reaction product with the test compound as compared to the amount of substrate or reaction product without the test compound identifies the test compound as an inhibitor of amylase activity.

The invention provides computer systems comprising a processor and a data storage device wherein said data storage device has stored thereon a polypeptide

sequence or a nucleic acid sequence of the invention (e.g., a polypeptide encoded by a nucleic acid of the invention). In one aspect, the computer system can further comprise a sequence comparison algorithm and a data storage device having at least one reference sequence stored thereon. In another aspect, the sequence comparison algorithm comprises a computer program that indicates polymorphisms. In one aspect, the computer system can further comprise an identifier that identifies one or more features in said sequence. The invention provides computer readable media having stored thereon a polypeptide sequence or a nucleic acid sequence of the invention. The invention provides methods for identifying a feature in a sequence comprising the steps of: (a) reading the sequence using a computer program which identifies one or more features in a sequence, wherein the sequence comprises a polypeptide sequence or a nucleic acid sequence of the invention; and (b) identifying one or more features in the sequence with the computer program. The invention provides methods for comparing a first sequence to a second sequence comprising the steps of: (a) reading the first sequence and the second sequence through use of a computer program which compares sequences, wherein the first sequence comprises a polypeptide sequence or a nucleic acid sequence of the invention; and (b) determining differences between the first sequence and the second sequence with the computer program. The step of determining differences between the first sequence and the second sequence can further comprise the step of identifying polymorphisms. In one aspect, the method can further comprise an identifier that identifies one or more features in a sequence. In another aspect, the method can comprise reading the first sequence using a computer program and identifying one or more features in the sequence.

The invention provides methods for isolating or recovering a nucleic acid encoding a polypeptide having an amylase activity from an environmental sample comprising the steps of: (a) providing an amplification primer sequence pair for amplifying a nucleic acid encoding a polypeptide having an amylase activity, wherein the primer pair is capable of amplifying a nucleic acid of the invention; (b) isolating a nucleic acid from the environmental sample or treating the environmental sample such that nucleic acid in the sample is accessible for hybridization to the amplification primer pair; and, (c) combining the nucleic acid of step (b) with the amplification primer pair of step (a) and amplifying nucleic acid from the environmental sample, thereby isolating or recovering a nucleic acid encoding a polypeptide having an amylase activity from an environmental sample. One or each member of the amplification primer sequence pair

can comprise an oligonucleotide comprising at least about 10 to 50 consecutive bases of a sequence of the invention.

The invention provides methods for isolating or recovering a nucleic acid encoding a polypeptide having an amylase activity from an environmental sample comprising the steps of: (a) providing a polynucleotide probe comprising a nucleic acid of the invention or a subsequence thereof; (b) isolating a nucleic acid from the environmental sample or treating the environmental sample such that nucleic acid in the sample is accessible for hybridization to a polynucleotide probe of step (a); (c) combining the isolated nucleic acid or the treated environmental sample of step (b) with the polynucleotide probe of step (a); and (d) isolating a nucleic acid that specifically hybridizes with the polynucleotide probe of step (a), thereby isolating or recovering a nucleic acid encoding a polypeptide having an amylase activity from an environmental sample. The environmental sample can comprise a water sample, a liquid sample, a soil sample, an air sample or a biological sample. In one aspect, the biological sample can be derived from a bacterial cell, a protozoan cell, an insect cell, a yeast cell, a plant cell, a fungal cell or a mammalian cell.

The invention provides methods of generating a variant of a nucleic acid encoding a polypeptide having an amylase activity comprising the steps of: (a) providing a template nucleic acid comprising a nucleic acid of the invention; and (b) modifying, deleting or adding one or more nucleotides in the template sequence, or a combination thereof, to generate a variant of the template nucleic acid. In one aspect, the method can further comprise expressing the variant nucleic acid to generate a variant amylase polypeptide. The modifications, additions or deletions can be introduced by a method comprising error-prone PCR, shuffling, oligonucleotide-directed mutagenesis, assembly PCR, sexual PCR mutagenesis, *in vivo* mutagenesis, cassette mutagenesis, recursive ensemble mutagenesis, exponential ensemble mutagenesis, site-specific mutagenesis, gene reassembly, gene site saturated mutagenesis (GSSM), synthetic ligation reassembly (SLR) or a combination thereof. In another aspect, the modifications, additions or deletions are introduced by a method comprising recombination, recursive sequence recombination, phosphothioate-modified DNA mutagenesis, uracil-containing template mutagenesis, gapped duplex mutagenesis, point mismatch repair mutagenesis, repair-deficient host strain mutagenesis, chemical mutagenesis, radiogenic mutagenesis, deletion mutagenesis, restriction-selection mutagenesis, restriction-purification mutagenesis,

artificial gene synthesis, ensemble mutagenesis, chimeric nucleic acid multimer creation and a combination thereof.

In one aspect, the method can be iteratively repeated until an amylase having an altered or different activity or an altered or different stability from that of a polypeptide encoded by the template nucleic acid is produced. In one aspect, the variant amylase polypeptide is thermotolerant, and retains some activity after being exposed to an elevated temperature. In another aspect, the variant amylase polypeptide has increased glycosylation as compared to the amylase encoded by a template nucleic acid. Alternatively, the variant amylase polypeptide has an amylase activity under a high temperature, wherein the amylase encoded by the template nucleic acid is not active under the high temperature. In one aspect, the method can be iteratively repeated until an amylase coding sequence having an altered codon usage from that of the template nucleic acid is produced. In another aspect, the method can be iteratively repeated until an amylase gene having higher or lower level of message expression or stability from that of the template nucleic acid is produced.

The invention provides methods for modifying codons in a nucleic acid encoding a polypeptide having an amylase activity to increase its expression in a host cell, the method comprising the following steps: (a) providing a nucleic acid of the invention encoding a polypeptide having an amylase activity; and, (b) identifying a non-preferred or a less preferred codon in the nucleic acid of step (a) and replacing it with a preferred or neutrally used codon encoding the same amino acid as the replaced codon, wherein a preferred codon is a codon over-represented in coding sequences in genes in the host cell and a non-preferred or less preferred codon is a codon under-represented in coding sequences in genes in the host cell, thereby modifying the nucleic acid to increase its expression in a host cell.

The invention provides methods for modifying codons in a nucleic acid encoding a polypeptide having an amylase activity; the method comprising the following steps: (a) providing a nucleic acid of the invention; and, (b) identifying a codon in the nucleic acid of step (a) and replacing it with a different codon encoding the same amino acid as the replaced codon, thereby modifying codons in a nucleic acid encoding an amylase.

The invention provides methods for modifying codons in a nucleic acid encoding a polypeptide having an amylase activity to increase its expression in a host

cell, the method comprising the following steps: (a) providing a nucleic acid of the invention encoding an amylase polypeptide; and, (b) identifying a non-preferred or a less preferred codon in the nucleic acid of step (a) and replacing it with a preferred or neutrally used codon encoding the same amino acid as the replaced codon, wherein a preferred codon is a codon over-represented in coding sequences in genes in the host cell and a non-preferred or less preferred codon is a codon under-represented in coding sequences in genes in the host cell, thereby modifying the nucleic acid to increase its expression in a host cell.

The invention provides methods for modifying a codon in a nucleic acid encoding a polypeptide having an amylase activity to decrease its expression in a host cell, the method comprising the following steps: (a) providing a nucleic acid of the invention; and (b) identifying at least one preferred codon in the nucleic acid of step (a) and replacing it with a non-preferred or less preferred codon encoding the same amino acid as the replaced codon, wherein a preferred codon is a codon over-represented in coding sequences in genes in a host cell and a non-preferred or less preferred codon is a codon under-represented in coding sequences in genes in the host cell, thereby modifying the nucleic acid to decrease its expression in a host cell. In one aspect, the host cell can be a bacterial cell, a fungal cell, an insect cell, a yeast cell, a plant cell or a mammalian cell.

The invention provides methods for producing a library of nucleic acids encoding a plurality of modified amylase active sites or substrate binding sites, wherein the modified active sites or substrate binding sites are derived from a first nucleic acid comprising a sequence encoding a first active site or a first substrate binding site the method comprising the following steps: (a) providing a first nucleic acid encoding a first active site or first substrate binding site, wherein the first nucleic acid sequence comprises a sequence that hybridizes under stringent conditions to a nucleic acid of the invention, and the nucleic acid encodes an amylase active site or an amylase substrate binding site; (b) providing a set of mutagenic oligonucleotides that encode naturally-occurring amino acid variants at a plurality of targeted codons in the first nucleic acid; and, (c) using the set of mutagenic oligonucleotides to generate a set of active site-encoding or substrate binding site-encoding variant nucleic acids encoding a range of amino acid variations at each amino acid codon that was mutagenized, thereby producing a library of nucleic acids encoding a plurality of modified amylase active sites or substrate binding sites. In one

aspect, the method comprises mutagenizing the first nucleic acid of step (a) by a method comprising an optimized directed evolution system, gene site-saturation mutagenesis (GSSM), synthetic ligation reassembly (SLR), error-prone PCR, shuffling, oligonucleotide-directed mutagenesis, assembly PCR, sexual PCR mutagenesis, in vivo mutagenesis, cassette mutagenesis, recursive ensemble mutagenesis, exponential ensemble mutagenesis, site-specific mutagenesis, gene reassembly, gene site saturated mutagenesis (GSSM), synthetic ligation reassembly (SLR) and a combination thereof. In another aspect, the method comprises mutagenizing the first nucleic acid of step (a) or variants by a method comprising recombination, recursive sequence recombination, phosphothioate-modified DNA mutagenesis, uracil-containing template mutagenesis, gapped duplex mutagenesis, point mismatch repair mutagenesis, repair-deficient host strain mutagenesis, chemical mutagenesis, radiogenic mutagenesis, deletion mutagenesis, restriction-selection mutagenesis, restriction-purification mutagenesis, artificial gene synthesis, ensemble mutagenesis, chimeric nucleic acid multimer creation and a combination thereof.

The invention provides methods for making a small molecule comprising the following steps: (a) providing a plurality of biosynthetic enzymes capable of synthesizing or modifying a small molecule, wherein one of the enzymes comprises an amylase enzyme encoded by a nucleic acid of the invention; (b) providing a substrate for at least one of the enzymes of step (a); and (c) reacting the substrate of step (b) with the enzymes under conditions that facilitate a plurality of biocatalytic reactions to generate a small molecule by a series of biocatalytic reactions. The invention provides methods for modifying a small molecule comprising the following steps: (a) providing an amylase enzyme, wherein the enzyme comprises a polypeptide of the invention, or, a polypeptide encoded by a nucleic acid of the invention, or a subsequence thereof; (b) providing a small molecule; and (c) reacting the enzyme of step (a) with the small molecule of step (b) under conditions that facilitate an enzymatic reaction catalyzed by the amylase enzyme, thereby modifying a small molecule by an amylase enzymatic reaction. In one aspect, the method can comprise a plurality of small molecule substrates for the enzyme of step (a), thereby generating a library of modified small molecules produced by at least one enzymatic reaction catalyzed by the amylase enzyme. In one aspect, the method can comprise a plurality of additional enzymes under conditions that facilitate a plurality of biocatalytic reactions by the enzymes to form a library of modified small molecules

produced by the plurality of enzymatic reactions. In another aspect, the method can further comprise the step of testing the library to determine if a particular modified small molecule which exhibits a desired activity is present within the library. The step of testing the library can further comprise the steps of systematically eliminating all but one of the biocatalytic reactions used to produce a portion of the plurality of the modified small molecules within the library by testing the portion of the modified small molecule for the presence or absence of the particular modified small molecule with a desired activity, and identifying at least one specific biocatalytic reaction that produces the particular modified small molecule of desired activity.

The invention provides methods for determining a functional fragment of an amylase enzyme comprising the steps of: (a) providing an amylase enzyme, wherein the enzyme comprises a polypeptide of the invention, or a polypeptide encoded by a nucleic acid of the invention, or a subsequence thereof; and (b) deleting a plurality of amino acid residues from the sequence of step (a) and testing the remaining subsequence for an amylase activity, thereby determining a functional fragment of an amylase enzyme. In one aspect, the amylase activity is measured by providing an amylase substrate and detecting a decrease in the amount of the substrate or an increase in the amount of a reaction product.

The invention provides methods for whole cell engineering of new or modified phenotypes by using real-time metabolic flux analysis, the method comprising the following steps: (a) making a modified cell by modifying the genetic composition of a cell, wherein the genetic composition is modified by addition to the cell of a nucleic acid of the invention; (b) culturing the modified cell to generate a plurality of modified cells; (c) measuring at least one metabolic parameter of the cell by monitoring the cell culture of step (b) in real time; and, (d) analyzing the data of step (c) to determine if the measured parameter differs from a comparable measurement in an unmodified cell under similar conditions, thereby identifying an engineered phenotype in the cell using real-time metabolic flux analysis. In one aspect, the genetic composition of the cell can be modified by a method comprising deletion of a sequence or modification of a sequence in the cell, or, knocking out the expression of a gene. In one aspect, the method can further comprise selecting a cell comprising a newly engineered phenotype. In another aspect, the method can comprise culturing the selected cell, thereby generating a new cell strain comprising a newly engineered phenotype.

The invention provides methods for hydrolyzing a starch comprising the following steps: (a) providing a polypeptide having an amylase activity, wherein the polypeptide comprises a polypeptide of the invention; (b) providing a composition comprising a starch; and (c) contacting the polypeptide of step (a) with the composition of step (b) under conditions wherein the polypeptide hydrolyzes the starch. In one aspect, the composition comprising starch that comprises an a-1,4-glucosidic bond or an a-1,6-glucosidic bond. In one aspect, the amylase activity is an a-amylase activity. In one aspect, the a-amylase activity hydrolyzes internal bonds in a starch or other polysaccharide.

The invention provides methods for liquefying or removing a starch from a composition comprising the following steps: (a) providing a polypeptide having an amylase activity, wherein the polypeptide comprises a polypeptide of the invention; (b) providing a composition comprising a starch; and (c) contacting the polypeptide of step (a) with the composition of step (b) under conditions wherein the polypeptide removes or liquefies the starch.

The invention provides methods of increasing thermotolerance or thermostability of an amylase polypeptide, the method comprising glycosylating an amylase polypeptide, wherein the polypeptide comprises at least thirty contiguous amino acids of a polypeptide of the invention; or a polypeptide encoded by a nucleic acid sequence of the invention, thereby increasing the thermotolerance or thermostability of the amylase polypeptide. In one aspect, the amylase specific activity can be thermostable or thermotolerant at a temperature in the range from greater than about 37°C to about 95°C.

The invention provides methods for overexpressing a recombinant amylase polypeptide in a cell comprising expressing a vector comprising a nucleic acid comprising a nucleic acid of the invention or a nucleic acid sequence of the invention, wherein the sequence identities are determined by analysis with a sequence comparison algorithm or by visual inspection, wherein overexpression is effected by use of a high activity promoter, a dicistronic vector or by gene amplification of the vector.

The invention provides detergent compositions comprising a polypeptide of the invention or a polypeptide encoded by a nucleic acid of the invention, wherein the polypeptide comprises an amylase activity. In one aspect, the amylase can be a

nonsurface-active amylase. In another aspect, the amylase can be a surface-active amylase.

The invention provides methods for washing an object comprising the following steps: (a) providing a composition comprising a polypeptide having an amylase activity, wherein the polypeptide comprises: a polypeptide of the invention or a polypeptide encoded by a nucleic acid of the invention; (b) providing an object; and (c) contacting the polypeptide of step (a) and the object of step (b) under conditions wherein the composition can wash the object.

The invention provides methods for hydrolyzing starch, e.g., in a feed or a food prior to consumption by an animal, comprising the following steps: (a) obtaining a composition, e.g., a feed material, comprising a starch, wherein the polypeptide comprises: a polypeptide of the invention or a polypeptide encoded by a nucleic acid of the invention; and (b) adding the polypeptide of step (a) to the composition, e.g., the feed or food material, in an amount sufficient for a sufficient time period to cause hydrolysis of the starch, thereby hydrolyzing the starch. In one aspect, the food or feed comprises rice, corn, barley, wheat, legumes, or potato.

The invention provides methods for textile desizing comprising the following steps: (a) providing a polypeptide having an amylase activity, wherein the polypeptide comprises a polypeptide of the invention or a polypeptide encoded by a nucleic acid of the invention; (b) providing a fabric; and (c) contacting the polypeptide of step (a) and the fabric of step (b) under conditions wherein the amylase can desize the fabric.

The invention provides methods for deinking of paper or fibers comprising the following steps: (a) providing a polypeptide having an amylase activity, wherein the polypeptide comprises a polypeptide of the invention; (b) providing a composition comprising paper or fiber; and (c) contacting the polypeptide of step (a) and the composition of step (b) under conditions wherein the polypeptide can deink the paper or fiber.

The invention provides methods for treatment of lignocellulosic fibers comprising the following steps: (a) providing a polypeptide having an amylase activity, wherein the polypeptide comprises a polypeptide of the invention; (b) providing a lignocellulosic fiber; and (c) contacting the polypeptide of step (a) and the fiber of step

(b) under conditions wherein the polypeptide can treat the fiber thereby improving the fiber properties.

The invention provides methods for producing a high-maltose or a high-glucose syrup comprising the following steps: (a) providing a polypeptide having an amylase activity, wherein the polypeptide comprises an enzyme of the invention; (b) providing a composition comprising a starch; and (c) contacting the polypeptide of step (a) and the fabric of step (b) under conditions wherein the polypeptide of step (a) can liquefy the composition of step (b) thereby producing a soluble starch hydrolysate and saccharify the soluble starch hydrolysate thereby producing the syrup. In one aspect, the starch can be from rice, corn, barley, wheat, legumes, potato, or sweet potato.

The invention provides methods for improving the flow of the starch-containing production fluids comprising the following steps: (a) providing a polypeptide having an amylase activity, wherein the polypeptide comprises a polypeptide of the invention; (b) providing production fluid; and (c) contacting the polypeptide of step (a) and the production fluid of step (b) under conditions wherein the amylase can hydrolyze the starch in the production fluid thereby improving its flow by decreasing its density. In one aspect, the production fluid can be from a subterranean formation.

The invention provides anti-staling compositions comprising a polypeptide of the invention or a polypeptide encoded by a nucleic acid of the invention. The invention provides methods for preventing staling of the baked products comprising the following steps: (a) providing a polypeptide having an amylase activity, wherein the polypeptide comprises a polypeptide of the invention; (b) providing a composition containing starch used for baking; (c) combining the polypeptide of step (a) with the composition of the step (b) under conditions wherein the polypeptide can hydrolyze the starch in the composition used for baking thereby preventing staling of the baked product. In one aspect, the baked product can be bread.

The invention provides methods for using amylase in brewing or alcohol production comprising the following steps: (a) providing a polypeptide having an amylase activity, wherein the polypeptide comprises a polypeptide of the invention; (b) providing a composition containing starch and used for brewing or in alcohol production; (c) combining the polypeptide of step (a) with the composition of the step (b) under conditions wherein the polypeptide can hydrolyze the starch in the composition used for

brewing or in alcohol production. In one aspect, the composition containing starch can be beer.

The invention provides methods of making a transgenic plant comprising the following steps: (a) introducing a heterologous nucleic acid sequence into the cell, wherein the heterologous nucleic sequence comprises a nucleic acid sequence of the invention, thereby producing a transformed plant cell; and (b) producing a transgenic plant from the transformed cell. In one aspect, the step (a) can further comprise introducing the heterologous nucleic acid sequence by electroporation or microinjection of plant cell protoplasts. In another aspect, the step (a) can further comprise introducing the heterologous nucleic acid sequence directly to plant tissue by DNA particle bombardment. Alternatively, the step (a) can further comprise introducing the heterologous nucleic acid sequence into the plant cell DNA using an Agrobacterium tumefaciens host. In one aspect, the plant cell can be a potato, corn, rice, wheat, tobacco, or barley cell.

The invention provides methods of expressing a heterologous nucleic acid sequence in a plant cell comprising the following steps: (a) transforming the plant cell with a heterologous nucleic acid sequence operably linked to a promoter, wherein the heterologous nucleic sequence comprises a nucleic acid of the invention; (b) growing the plant under conditions wherein the heterologous nucleic acids sequence is expressed in the plant cell.

The invention also provides a process for preparing a dough or a baked product prepared from the dough which comprises adding an amylase of the invention to the dough in an amount which is effective to retard the staling of the bread. The invention also provides a dough comprising said amylase and a premix comprising flour together with said amylase. Finally, the invention provides an enzymatic baking additive, which contains said amylase. The use of the amylase in accordance with the present invention provides an improved anti-staling effect as measured by, e.g. less crumb firming, retained crumb elasticity, improved slice-ability (e.g. fewer crumbs, non-gummy crumb), improved palatability or flavor.

The invention provides delayed release ("controlled release") compositions comprising an desired ingredient coated by a latex polymer (or equivalent) coating. In one aspect, the desired ingredient comprises an enzyme, e.g., an enzyme of the invention. In one aspect, the desired ingredient comprises a small molecule, a drug, a

polysaccharide, a lipid, a nucleic acid, a vitamin, an antibiotic or an insecticide. In one aspect, the desired ingredient comprises a pellet or a matrix, e. g. , a pellet or a matrix comprising an edible material (e. g. , as an animal food or feed or supplement or medicament). The invention also provides methods for the "controlled release" or "delayed release" of a composition, wherein the composition is coated by a latex polymer (or equivalent) coating.

In one aspect, the latex polymer coating comprises a latex paint, or equivalent. The latex polymer coating can comprise a (meth) acrylate, a vinyl acetate, a styrene, an ethylene, a vinyl chloride, a butadiene, a vinylidene chloride, a vinyl versatate, a vinyl propionate, a t-butyl acrylate, an acrylonitrile, a neoprene, a maleate, a fumarate, equivalents thereof, combinations thereof and/or derivatives thereof.

The details of one or more embodiments of the invention are set forth in the accompanying drawings and the description below. Other features, aspects, and advantages of the invention will be apparent from the description and drawings, and from the claims.

All publications, patents, patent applications, GenBank sequences and ATCC deposits, cited herein are hereby expressly incorporated by reference for all purposes.

As used herein, except where the context requires otherwise, the term "comprise" and variations of the term, such as "comprising", "comprises" and "comprised", are not intended to exclude other additives, components, integers or steps.

Reference to any prior art in the specification is not, and should not be taken as, an acknowledgment, or any form of suggestion, that this prior art forms part of the common general knowledge in Australia or any other jurisdiction or that this prior art could reasonably be expected to be ascertained, understood and regarded as relevant by a person skilled in the art.

DESCRIPTION OF DRAWINGS

Figure 1 is a block diagram of a computer system.

2013201807 18 Mar 2013

Figure 2 is a flow diagram illustrating one aspect of a process for comparing a new nucleotide or protein sequence with a database of sequences in order to determine the homology levels between the new sequence and the sequences in the database.

Figure 3 is a flow diagram illustrating one aspect of a process in a computer for determining whether two sequences are homologous.

Figure 4 is a flow diagram illustrating one aspect of an identifier process 300 for detecting the presence of a feature in a sequence.

Figure 5 is a graph showing the Residual activity of various amylases following heating to 90°C for 10 min in Example 1.

Figure 6 is a graph showing the net percent starch removed versus enzyme concentration in ADW wash test with bleach and chelators.

Figure 7 is a graph showing the activity of parental amylases at pH 8, 40°C in ADW formulation at 55°C.

Figure 8 is a graph of data regarding the H_2O_2 tolerance of the novel enzymes in Example 4.

Figure 9 is a graph of the pH and temperature data for a selection of the amylases characterized. Figure 9a shows the data at pH 8 and 40°C and Figure 9b shows the data at pH 10 and 50°C.

Figure 10 sets forth the sequences to be used in reassembly experiments with the enzymes.

Figure 11 illustrates a sample Standard Curve of the assay of Example 5.

Figure 12 illustrates the pH rate profiles for SEQ ID NO.: 127, which has a neutral optimum pH and SEQ ID NO.: 211, which has an optimum around pH 10.

Figure 13 shows the stability of exemplary amylases vs. a commercial enzyme, as discussed in Example 2.

Figure 14 shows the sequence alignments of hypothermophilic α-amylases, as set forth in Example 8. Figure 14a shows an alignment of amylase sequences. SEQ ID NO.: 81= an environmental clone; pyro = *Pyrococcus sp.* (strain:KOD1), Tachibana (1996) J. Ferment. Bioeng. 82:224-232; pyro2 = *Pyrococcus furiosus*, Appl. Environ. Microbiol. 63 (9):3569-3576, 1997; Thermo = *Thermococcus sp.*; Thermo2 = *Thermococcus hydrothermalis,* Leveque,E. *et al.* Patent: France 98.05655 05-MAY-1998. Figure 14b shows the amino acid sequence alignment of identified sequences: SEQ ID NO.: 81; pyro; SEQ ID NO.:75; SEQ ID NO.: 77; SEQ ID NO.: 83; SEQ ID NO.: 85; thermo2; SEQ ID NO.: 79 ; thermo ; pyro2 ; clone A; thermo3. Figure 14c shows the nucleic acid sequence alignment corresponding to the polypeptide sequence of Figures 5 and 6. SEQ ID NO.: 81; SEQ ID NO.:75; SEQ ID NO.: 77; SEQ ID NO.: 83; SEQ ID NO.: 85; SEQ ID NO.: 79; clone A; and SEQ ID NO.: 73.

Figure 15 is a neighbor-joining tree for *Thermococcales*.

Figure 16 shows sequences of exemplary sequences of the invention.

Figure 17 illustrates methods of the invention for liquefaction saccharification of starch, as described in detail, below.

Figure 18 illustrates Table 7, which lists the relative percent identities of exemplary sequences of the invention, as described in Example 8, below.

Figure 19 shows the pH profile of tested amylases of the invention and a commercial benchmark enzyme, as described in Example 15, below.

Figure 20 shows the temperature activity profiles of exemplary amylases of the invention, as described in Example 15, below.

Figure 21 shows enzyme activity (of exemplary amylases of the invention) in the presence of EDTA, as described in Example 15, below.

Figure 22 shows enzyme activity (of exemplary amylases of the invention) in the presence of peroxide hydroxide, as described in Example 15, below.

Figure 23 shows enzyme activity (of exemplary amylases of the invention) in the ADW solution (distilled water, hardening solution, bleach, chelators, surfactants) with soluble substrate (BODIPY-starch), as described in Example 15, below.

Figure 24 shows the results of the wash tests with starch-coated slides using exemplary amylases of the invention, as described in Example 15, below.

Figure 25 illustrates an exemplary corn wet milling process of the invention (using at least one enzyme of the invention).

Figure 26, Figure 27 and Figure 28 illustrate alternative exemplary starch processes, including starch liquefaction processes, of the invention (using at least one enzyme of the invention), as described in detail, below.

Figure 29 shows data summarizing these findings comparing amylase SEQ ID NO:437 with TERMAMYL™ SC (Novozymes A/S, Denmark) amylase in dry mill ethanol processing, as described in Example 1, below.

Figure 30 illustrates a pH activity profile of an exemplary enzyme of the invention (SEQ ID NO:594) in acetate buffer and phosphate buffer to determine the relative rate for the glucoamylase at each pH, as discussed in detail in Example 16, below.

Figure 31 illustrates a temperature activity profile of an exemplary enzyme of the invention (SEQ ID NO:594) in acetate buffer, as discussed in detail in Example 16, below.

Figure 32 illustrates a temperature stability profile of an exemplary enzyme of the invention (SEQ ID NO:594), as discussed in detail in Example 16, below.

Figure 33 illustrates a substrate utilization activity profile of an exemplary enzyme of the invention (SEQ ID NO:594) using the dextrins maltose (G2), maltotriose

(G3), panose (Pan), maltotetraose (G4), and maltoheptaose (G7), as discussed in detail in Example 16, below.

Figure 34 illustrates an exemplary glucoamylase-encoding nucleic acid of the invention, the genomic sequence set forth in SEQ ID NO:587. Coding sequences (exons) are denoted with the single-letter amino acid below it. Intron sequences are underlined.

Figure 35 is a chart describing selected characteristics of exemplary nucleic acids and polypeptides of the invention, as described in further detail, below.

Like reference symbols in the various drawings indicate like elements.

DETAILED DESCRIPTION

The invention provides amylase enzymes, e.g., an alpha amylases, polynucleotides encoding the enzymes, methods of making and using these polynucleotides and polypeptides. The invention is directed to novel polypeptides having an amylase activity, e.g., an alpha amylase activity, nucleic acids encoding them and antibodies that bind to them. The polypeptides of the invention can be used in a variety of diagnostic, therapeutic, and industrial contexts. The polypeptides of the invention can be used as, e.g., an additive for a detergent, for processing foods and for chemical synthesis utilizing a reverse reaction. Additionally, the polypeptides of the invention can be used in fabric treatment, alcohol production, and as additives to food or animal feed.

In one aspect, the amylases of the invention are active at a high and/or at a low temperature, or, over a wide range of temperature. For example, they can be active in the temperatures ranging between 20°C to 90°C, between 30°C to 80°C, or between 40°C to 70°C. The invention also provides amylases that have activity at alkaline pHs or at acidic pHs, e.g., low water acidity. In alternative aspects, the amylases of the invention can have activity in acidic pHs as low as pH 5.0, pH 4.5, pH 4.0, and pH 3.5. In alternative aspects, the amylases of the invention can have activity in alkaline pHs as high as pH 9.5, pH 10, pH 10.5, and pH 11. In one aspect, the amylases of the invention are active in the temperature range of between about 40°C to about 70°C under conditions of low water activity (low water content).

The invention also provides methods for further modifying the exemplary amylases of the invention to generate proteins with desirable properties. For example,

amylases generated by the methods of the invention can have altered enzymatic activity, thermal stability, pH/activity profile, pH/stability profile (such as increased stability at low, e.g. pH<6 or pH<5, or high, e.g. pH>9, pH values), stability towards oxidation, Ca^{2+} dependency, specific activity and the like. The invention provides for altering any property of interest. For instance, the alteration may result in a variant which, as compared to a parent enzyme, has altered enzymatic activity, or, pH or temperature activity profiles.

Definitions

The term "amylase" includes all polypeptides, e.g., enzymes, which catalyze the hydrolysis of a polysaccharide, e.g., a starch. The term "amylase" includes polypeptides having an a-amylase activity, a ß-amylase activity, a glucoamylase activity, a 1,4-a-D-glucan glucohydrolase activity, an exoamylase activity, a glucan a-maltotetrahydrolase activity, a maltase activity, an isomaltase activity, a glucan 1, 4, a-glucosidase activity, an a-glucosidase activity, a sucrase activity or an agarase activity (e.g., a ß-agarase activity). For example, an amylase activity of the invention includes a-amylase activity, including the ability to hydrolyze internal alpha-1,4-glucosidic linkages in starch to produce smaller molecular weight malto-dextrins. In one aspect, the a-amylase activity includes hydrolyzing internal alpha-1,4-glucosidic linkages in starch at random. An amylase activity of the invention includes polypeptides having glucoamylase activity, such as the ability to hydrolase glucose polymers linked by a-1,4- and a-1,6-glucosidic bonds. In one aspect, the polypeptides of the invention have glucoamylase activity, hydrolyzing internal a-1,4-glucosidic linkages to yield smaller molecular weight malto-dextrins. An amylase activity of the invention also includes glucan 1,4-a-glucosidase activity, or, 1,4-a-D-glucan glucohydrolase, commonly called glucoamylase but also called amyloglucosidase and ?-amylase that, in one aspect, releases ß-D-glucose from 1,4-a-, 1,6-a- and 1,3-a-linked glucans. An amylase activity of the invention also includes exo-amylase activity.

In one aspect, the glucoamylase activity comprises catalysis of the hydrolysis of glucosidic bonds. The glucoamylase activity can comprise catalyzing the step-wise hydrolytic release of D-glucose from the non-reducing ends of starch or other related dextrins. The glucoamylase activity can comprise a 1,4-a-D-glucan glucohydralase activity. The glucoamylase activity can comprise catalysis of the

hydrolysis of malto-dextrins resulting in the generation of free glucose. The glucoamylase activity can comprise an exoamylase activity. The glucoamylase activity can comprise an a-amylase or a ß-amylase activity. The hydrolyzed glucosidic bonds can comprise a-1,4-glucosidic bonds or a-1,6-glucosidic bonds. The glucoamylase activity can comprise hydrolyzing glucosidic bonds in a starch. The glucoamylase activity can further comprise hydrolyzing glucosidic bonds in the starch to produce maltodextrines. The glucoamylase activity can comprise cleaving a maltose or a D-glucose unit from non-reducing end of the starch.

An amylase activity of the invention also includes hydrolyzing a polysaccharide, e.g., a starch, at high temperatures, low temperatures, alkaline pHs and at acidic pHs. For example, in one aspect, the invention provides polypeptides, and nucleic acids encoding them, having an amylase, e.g., a glucoamylase, activity which is thermostable. The polypeptide can retain an amylase activity under conditions comprising a temperature range of between about 37°C to about 95°C; between about 55°C to about 85°C, between about 70°C to about 95°C, or, between about 90°C to about 95°C. In another aspect, a polypeptide of the invention can have a glucoamylase activity which is thermotolerant. The polypeptide can retain an amylase, e.g., a glucoamylase, activity after exposure to a temperature in the range from greater than 37°C to about 95°C or anywhere in the range from greater than 55°C to about 85°C. In one aspect, the polypeptide retains an amylase activity after exposure to a temperature in the range from greater than 90°C to about 95°C at pH 4.5.

An "amylase variant" comprises an amino acid sequence which is derived from the amino acid sequence of a "precursor amylase". The precursor amylase can include naturally-occurring amylases and recombinant amylases. The amino acid sequence of the amylase variant can be "derived" from the precursor amylase amino acid sequence by the substitution, deletion or insertion of one or more amino acids of the precursor amino acid sequence. Such modification can be of the "precursor DNA sequence" which encodes the amino acid sequence of the precursor amylase rather than manipulation of the precursor amylase enzyme per se. Suitable methods for such manipulation of the precursor DNA sequence include methods disclosed herein, as well as methods known to those skilled in the art.

The term "antibody" includes a peptide or polypeptide derived from, modeled after or substantially encoded by an immunoglobulin gene or immunoglobulin

genes, or fragments thereof, capable of specifically binding an antigen or epitope, see, e.g. Fundamental Immunology, Third Edition, W.E. Paul, ed., Raven Press, N.Y. (1993); Wilson (1994) J. Immunol. Methods 175:267-273; Yarmush (1992) J. Biochem. Biophys. Methods 25:85-97. The term antibody includes antigen-binding portions, i.e., "antigen binding sites," (e.g., fragments, subsequences, complementarity determining regions (CDRs)) that retain capacity to bind antigen, including (i) a Fab fragment, a monovalent fragment consisting of the VL, VH, CL and CH1 domains; (ii) a F(ab')2 fragment, a bivalent fragment comprising two Fab fragments linked by a disulfide bridge at the hinge region; (iii) a Fd fragment consisting of the VH and CH1 domains; (iv) a Fv fragment consisting of the VL and VH domains of a single arm of an antibody, (v) a dAb fragment (Ward et al., (1989) Nature 341:544-546), which consists of a VH domain; and (vi) an isolated complementarity determining region (CDR). Single chain antibodies are also included by reference in the term "antibody."

The terms "array" or "microarray" or "biochip" or "chip" as used herein is a plurality of target elements, each target element comprising a defined amount of one or more polypeptides (including antibodies) or nucleic acids immobilized onto a defined area of a substrate surface, as discussed in further detail, below.

As used herein, the terms "computer," "computer program" and "processor" are used in their broadest general contexts and incorporate all such devices, as described in detail, below. A "coding sequence of" or a "sequence encodes" a particular polypeptide or protein, is a nucleic acid sequence which is transcribed and translated into a polypeptide or protein when placed under the control of appropriate regulatory sequences.

The term "expression cassette" as used herein refers to a nucleotide sequence which is capable of affecting expression of a structural gene (i.e., a protein coding sequence, such as an amylase of the invention) in a host compatible with such sequences. Expression cassettes include at least a promoter operably linked with the polypeptide coding sequence; and, optionally, with other sequences, e.g., transcription termination signals. Additional factors necessary or helpful in effecting expression may also be used, e.g., enhancers. Thus, expression cassettes also include plasmids, expression vectors, recombinant viruses, any form of recombinant "naked DNA" vector, and the like.

"Operably linked" as used herein refers to a functional relationship between two or more nucleic acid (e.g., DNA) segments. Typically, it refers to the functional relationship of transcriptional regulatory sequence to a transcribed sequence. For example, a promoter is operably linked to a coding sequence, such as a nucleic acid of the invention, if it stimulates or modulates the transcription of the coding sequence in an appropriate host cell or other expression system. Generally, promoter transcriptional regulatory sequences that are operably linked to a transcribed sequence are physically contiguous to the transcribed sequence, i.e., they are *cis*-acting. However, some transcriptional regulatory sequences, such as enhancers, need not be physically contiguous or located in close proximity to the coding sequences whose transcription they enhance.

A "vector" comprises a nucleic acid which can infect, transfect, transiently or permanently transduce a cell. It will be recognized that a vector can be a naked nucleic acid, or a nucleic acid complexed with protein or lipid. The vector optionally comprises viral or bacterial nucleic acids and/or proteins, and/or membranes (e.g., a cell membrane, a viral lipid envelope, etc.). Vectors include, but are not limited to replicons (e.g., RNA replicons, bacteriophages) to which fragments of DNA may be attached and become replicated. Vectors thus include, but are not limited to RNA, autonomous self-replicating circular or linear DNA or RNA (e.g., plasmids, viruses, and the like, see, e.g., U.S. Patent No. 5,217,879), and include both the expression and non-expression plasmids. Where a recombinant microorganism or cell culture is described as hosting an "expression vector" this includes both extra-chromosomal circular and linear DNA and DNA that has been incorporated into the host chromosome(s). Where a vector is being maintained by a host cell, the vector may either be stably replicated by the cells during mitosis as an autonomous structure, or is incorporated within the host's genome.

As used herein, the term "promoter" includes all sequences capable of driving transcription of a coding sequence in a cell, e.g., a plant cell. Thus, promoters used in the constructs of the invention include *cis*-acting transcriptional control elements and regulatory sequences that are involved in regulating or modulating the timing and/or rate of transcription of a gene. For example, a promoter can be a *cis*-acting transcriptional control element, including an enhancer, a promoter, a transcription terminator, an origin of replication, a chromosomal integration sequence, 5' and 3' untranslated regions, or an intronic sequence, which are involved in transcriptional

regulation. These cis-acting sequences typically interact with proteins or other biomolecules to carry out (turn on/off, regulate, modulate, etc.) transcription. "Constitutive" promoters are those that drive expression continuously under most environmental conditions and states of development or cell differentiation. "Inducible" or "regulatable" promoters direct expression of the nucleic acid of the invention under the influence of environmental conditions or developmental conditions. Examples of environmental conditions that may affect transcription by inducible promoters include anaerobic conditions, elevated temperature, drought, or the presence of light.

"Tissue-specific" promoters are transcriptional control elements that are only active in particular cells or tissues or organs, e.g., in plants or animals. Tissue-specific regulation may be achieved by certain intrinsic factors which ensure that genes encoding proteins specific to a given tissue are expressed. Such factors are known to exist in mammals and plants so as to allow for specific tissues to develop.

The term "plant" includes whole plants, plant parts (e.g., leaves, stems, flowers, roots, etc.), plant protoplasts, seeds and plant cells and progeny of same. The class of plants which can be used in the method of the invention is generally as broad as the class of higher plants amenable to transformation techniques, including angiosperms (monocotyledonous and dicotyledonous plants), as well as gymnosperms. It includes plants of a variety of ploidy levels, including polyploid, diploid, haploid and hemizygous states. As used herein, the term "transgenic plant" includes plants or plant cells into which a heterologous nucleic acid sequence has been inserted, e.g., the nucleic acids and various recombinant constructs (e.g., expression cassettes) of the invention.

"Plasmids" can be commercially available, publicly available on an unrestricted basis, or can be constructed from available plasmids in accord with published procedures. Equivalent plasmids to those described herein are known in the art and will be apparent to the ordinarily skilled artisan.

The term "gene" includes a nucleic acid sequence comprising a segment of DNA involved in producing a transcription product (e.g., a message), which in turn is translated to produce a polypeptide chain, or regulates gene transcription, reproduction or stability. Genes can include regions preceding and following the coding region, such as leader and trailer, promoters and enhancers, as well as, where applicable, intervening sequences (introns) between individual coding segments (exons).

The phrases "nucleic acid" or "nucleic acid sequence" includes oligonucleotide, nucleotide, polynucleotide, or to a fragment of any of these, to DNA or RNA (e.g., mRNA, rRNA, tRNA) of genomic or synthetic origin which may be single-stranded or double-stranded and may represent a sense or antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material, natural or synthetic in origin, including, e.g., iRNA, ribonucleoproteins (e.g., iRNPs). The term encompasses nucleic acids, i.e., oligonucleotides, containing known analogues of natural nucleotides. The term also encompasses nucleic-acid-like structures with synthetic backbones, see e.g., Mata (1997) Toxicol. Appl. Pharmacol. 144:189-197; Strauss-Soukup (1997) Biochemistry 36:8692-8698; Samstag (1996) Antisense Nucleic Acid Drug Dev 6:153-156.

"Amino acid" or "amino acid sequence" include an oligopeptide, peptide, polypeptide, or protein sequence, or to a fragment, portion, or subunit of any of these, and to naturally occurring or synthetic molecules. The terms "polypeptide" and "protein" include amino acids joined to each other by peptide bonds or modified peptide bonds, i.e., peptide isosteres, and may contain modified amino acids other than the 20 gene-encoded amino acids. The term "polypeptide" also includes peptides and polypeptide fragments, motifs and the like. The term also includes glycosylated polypeptides. The peptides and polypeptides of the invention also include all "mimetic" and "peptidomimetic" forms, as described in further detail, below.

The term "isolated" includes a material removed from its original environment, e.g., the natural environment if it is naturally occurring. For example, a naturally occurring polynucleotide or polypeptide present in a living animal is not isolated, but the same polynucleotide or polypeptide, separated from some or all of the coexisting materials in the natural system, is isolated. Such polynucleotides could be part of a vector and/or such polynucleotides or polypeptides could be part of a composition, and still be isolated in that such vector or composition is not part of its natural environment. As used herein, an isolated material or composition can also be a "purified" composition, i.e., it does not require absolute purity; rather, it is intended as a relative definition. Individual nucleic acids obtained from a library can be conventionally purified to electrophoretic homogeneity. In alternative aspects, the invention provides nucleic acids which have been purified from genomic DNA or from other sequences in a

library or other environment by at least one, two, three, four, five or more orders of magnitude.

As used herein, the term "recombinant" can include nucleic acids adjacent to a "backbone" nucleic acid to which it is not adjacent in its natural environment. In one aspect, nucleic acids represent 5% or more of the number of nucleic acid inserts in a population of nucleic acid "backbone molecules." "Backbone molecules" according to the invention include nucleic acids such as expression vectors, self-replicating nucleic acids, viruses, integrating nucleic acids, and other vectors or nucleic acids used to maintain or manipulate a nucleic acid insert of interest. In one aspect, the enriched nucleic acids represent 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more of the number of nucleic acid inserts in the population of recombinant backbone molecules. "Recombinant" polypeptides or proteins refer to polypeptides or proteins produced by recombinant DNA techniques; e.g., produced from cells transformed by an exogenous DNA construct encoding the desired polypeptide or protein. "Synthetic" polypeptides or protein are those prepared by chemical synthesis, as described in further detail, below.

A promoter sequence can be "operably linked to" a coding sequence when RNA polymerase which initiates transcription at the promoter will transcribe the coding sequence into mRNA, as discussed further, below.

"Oligonucleotide" includes either a single stranded polydeoxynucleotide or two complementary polydeoxynucleotide strands which may be chemically synthesized. Such synthetic oligonucleotides have no 5' phosphate and thus will not ligate to another oligonucleotide without adding a phosphate with an ATP in the presence of a kinase. A synthetic oligonucleotide can ligate to a fragment that has not been dephosphorylated.

The phrase "substantially identical" in the context of two nucleic acids or polypeptides, can refer to two or more sequences that have, e.g., at least about 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more nucleotide or amino acid residue (sequence) identity, when

compared and aligned for maximum correspondence, as measured using one any known sequence comparison algorithm, as discussed in detail below, or by visual inspection. In alternative aspects, the invention provides nucleic acid and polypeptide sequences having substantial identity to an exemplary sequence of the invention over a region of at least about 10, 20, 30, 40, 50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000 or more residues, or a region ranging from between about 50 residues to the full length of the nucleic acid or polypeptide. Nucleic acid sequences of the invention can be substantially identical over the entire length of a polypeptide coding region.

A "substantially identical" amino acid sequence also can include a sequence that differs from a reference sequence by one or more conservative or non-conservative amino acid substitutions, deletions, or insertions, particularly when such a substitution occurs at a site that is not the active site of the molecule, and provided that the polypeptide essentially retains its functional properties. A conservative amino acid substitution, for example, substitutes one amino acid for another of the same class (e.g., substitution of one hydrophobic amino acid, such as isoleucine, valine, leucine, or methionine, for another, or substitution of one polar amino acid for another, such as substitution of arginine for lysine, glutamic acid for aspartic acid or glutamine for asparagine). One or more amino acids can be deleted, for example, from an amylase, resulting in modification of the structure of the polypeptide, without significantly altering its biological activity. For example, amino- or carboxyl-terminal amino acids that are not required for amylase activity can be removed.

"Hybridization" includes the process by which a nucleic acid strand joins with a complementary strand through base pairing. Hybridization reactions can be sensitive and selective so that a particular sequence of interest can be identified even in samples in which it is present at low concentrations. Stringent conditions can be defined by, for example, the concentrations of salt or formamide in the prehybridization and hybridization solutions, or by the hybridization temperature, and are well known in the art. For example, stringency can be increased by reducing the concentration of salt, increasing the concentration of formamide, or raising the hybridization temperature, altering the time of hybridization, as described in detail, below. In alternative aspects, nucleic acids of the invention are defined by their ability to hybridize under various stringency conditions (e.g., high, medium, and low), as set forth herein.

"Variant" includes polynucleotides or polypeptides of the invention modified at one or more base pairs, codons, introns, exons, or amino acid residues (respectively) yet still retain the biological activity of an amylase of the invention. Variants can be produced by any number of means included methods such as, for example, error-prone PCR, shuffling, oligonucleotide-directed mutagenesis, assembly PCR, sexual PCR mutagenesis, in vivo mutagenesis, cassette mutagenesis, recursive ensemble mutagenesis, exponential ensemble mutagenesis, site-specific mutagenesis, gene reassembly, GSSM and any combination thereof. Techniques for producing variant amylase having activity at a pH or temperature, for example, that is different from a wild-type amylase, are included herein.

The term "saturation mutagenesis" or "GSSM" includes a method that uses degenerate oligonucleotide primers to introduce point mutations into a polynucleotide, as described in detail, below.

The term "optimized directed evolution system" or "optimized directed evolution" includes a method for reassembling fragments of related nucleic acid sequences, e.g., related genes, and explained in detail, below.

The term "synthetic ligation reassembly" or "SLR" includes a method of ligating oligonucleotide fragments in a non-stochastic fashion, and explained in detail, below.

Generating and Manipulating Nucleic Acids

In one aspect, the invention provides isolated or recombinant nucleic acids comprising a nucleic acid sequence having at least about 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more, or complete (100%) sequence identity to an exemplary nucleic acid of the invention over a region of at least about 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400, 1450, 1500, 1550 or more, residues. In one aspect, the nucleic acid encodes at least one polypeptide having an amylase activity, e.g., an alpha amylase activity.

For example, the following table describes some exemplary amylase-encoding nucleic acids of the invention, e.g., the invention provides an amylase having a sequence as set forth in SEQ ID NO:474, having an exemplary coding sequence as set forth in SEQ ID NO:473, and in one aspect is encoded by a gene, including introns and exons, having a sequence as set forth in SEQ ID NO:467 (including exons having sequences as set forth in SEQ ID NO:468, SEQ ID NO:469, SEQ ID NO:470, SEQ ID NO:471 and SEQ ID NO:472); etc.:

Amylase	SEQ ID NO: of full gene (exons and introns)	SEQ ID NOS: of exon sequences	SEQ ID NO: of DNA sequence of coding sequence (exons only)	SEQ ID NO: of protein sequence of coding sequence (exons only)	TOTAL
A	460, 461	N/A	460	461	460, 461
B	462	463, 464	465	466	462-466
C	467	468-472	473	474	467-474
D	475	476-477	478	479	475-479
E	480	481-483	484	485	480-485
F	486	487-491	492	493	486-493
G	494	495-497	498	499	494-499
H	500	501-508	509	510	500-510
I	511	512-514	515	516	511-516
J	517, 518	N/A	517	518	517, 518
K	519	520-521	522	523	519-523
L	524	525-526	527	528	524-528
M	529	530-531	532	533	529-533
N	534	535-538	539	540	534-540
O	541	542-543	544	545	541-545
P	546	547-551	552	553	546-553
Q	554	555-557	558	559	554-559
R	560	561-564	565	566	560-566
S	587	588-592	593	594	587-594

The above listed amylases (described as A thru S) and the nucleic acids that encode them have a common novelty in that they were initially isolated/ derived from fungal sources.

The invention also provides glucoamylases, such as the enzyme having a sequence as set forth in SEQ ID NO:594 encoded by the 4111 residues of the genomic SEQ ID NO:587, or, the 1854 residue long cDNA of SEQ ID NO:593). The genomic SEQ ID NO:587, comprises introns and exons, and the exons can be described as encoding polypeptide fragments having a sequence as set forth in SEQ ID NO:588, SEQ

ID NO:589, SEQ ID NO:590, SEQ ID NO:591, SEQ ID NO:592. In one aspect, the "mature" processed glucoamylase consisting of residues 32 to 617 of SEQ ID NO: 594.

The invention provides isolated and recombinant nucleic acids, including expression cassettes such as expression vectors encoding the polypeptides of the invention. The invention provides probes comprising or consisting of nucleic acids of the invention. The invention also includes methods for discovering new amylase sequences using the nucleic acids of the invention. The invention also includes methods for inhibiting the expression of amylase genes, transcripts and polypeptides using the nucleic acids of the invention. Also provided are methods for modifying the nucleic acids of the invention by, e.g., synthetic ligation reassembly, optimized directed evolution system and/or gene site saturation mutagenesis (GSSM™).

The nucleic acids of the invention can be made, isolated and/or manipulated by, e.g., cloning and expression of cDNA libraries, amplification of message or genomic DNA by PCR, and the like. In practicing the methods of the invention, homologous genes can be modified by manipulating a template nucleic acid, as described herein. The invention can be practiced in conjunction with any method or protocol or device known in the art, which are well described in the scientific and patent literature.

General Techniques

The nucleic acids used to practice this invention, whether RNA, iRNA, antisense nucleic acid, cDNA, genomic DNA, vectors, viruses or hybrids thereof, may be isolated from a variety of sources, genetically engineered, amplified, and/or expressed/ generated recombinantly. Recombinant polypeptides generated from these nucleic acids can be individually isolated or cloned and tested for a desired activity. Any recombinant expression system can be used, including bacterial, mammalian, yeast, insect or plant cell expression systems.

Alternatively, these nucleic acids can be synthesized *in vitro* by well-known chemical synthesis techniques, as described in, e.g., Adams (1983) J. Am. Chem. Soc. 105:661; Belousov (1997) Nucleic Acids Res. 25:3440-3444; Frenkel (1995) Free Radic. Biol. Med. 19:373-380; Blommers (1994) Biochemistry 33:7886-7896; Narang (1979) Meth. Enzymol. 68:90; Brown (1979) Meth. Enzymol. 68:109; Beaucage (1981) Tetra. Lett. 22:1859; U.S. Patent No. 4,458,066.

Techniques for the manipulation of nucleic acids, such as, e.g., subcloning, labeling probes (e.g., random-primer labeling using Klenow polymerase, nick translation, amplification), sequencing, hybridization and the like are well described in the scientific and patent literature, see, e.g., Sambrook, ed., MOLECULAR CLONING: A LABORATORY MANUAL (2ND ED.), Vols. 1-3, Cold Spring Harbor Laboratory, (1989); CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, Ausubel, ed. John Wiley & Sons, Inc., New York (1997); LABORATORY TECHNIQUES IN BIOCHEMISTRY AND MOLECULAR BIOLOGY: HYBRIDIZATION WITH NUCLEIC ACID PROBES, Part I. Theory and Nucleic Acid Preparation, Tijssen, ed. Elsevier, N.Y. (1993).

Another useful means of obtaining and manipulating nucleic acids used to practice the methods of the invention is to clone from genomic samples, and, if desired, screen and re-clone inserts isolated or amplified from, e.g., genomic clones or cDNA clones. Sources of nucleic acid used in the methods of the invention include genomic or cDNA libraries contained in, e.g., mammalian artificial chromosomes (MACs), see, e.g., U.S. Patent Nos. 5,721,118; 6,025,155; human artificial chromosomes, see, e.g., Rosenfeld (1997) Nat. Genet. 15:333-335; yeast artificial chromosomes (YAC); bacterial artificial chromosomes (BAC); P1 artificial chromosomes, see, e.g., Woon (1998) Genomics 50:306-316; P1-derived vectors (PACs), see, e.g., Kern (1997) Biotechniques 23:120-124; cosmids, recombinant viruses, phages or plasmids.

In one aspect, a nucleic acid encoding a polypeptide of the invention is assembled in appropriate phase with a leader sequence capable of directing secretion of the translated polypeptide or fragment thereof.

The invention provides fusion proteins and nucleic acids encoding them. A polypeptide of the invention can be fused to a heterologous peptide or polypeptide, such as N-terminal identification peptides which impart desired characteristics, such as increased stability or simplified purification. Peptides and polypeptides of the invention can also be synthesized and expressed as fusion proteins with one or more additional domains linked thereto for, e.g., producing a more immunogenic peptide, to more readily isolate a recombinantly synthesized peptide, to identify and isolate antibodies and antibody-expressing B cells, and the like. Detection and purification facilitating domains include, e.g., metal chelating peptides such as polyhistidine tracts and histidine-tryptophan modules that allow purification on immobilized metals, protein A domains

that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAGS extension/affinity purification system (Immunex Corp, Seattle WA). The inclusion of a cleavable linker sequences such as Factor Xa or enterokinase (Invitrogen, San Diego CA) between a purification domain and the motif-comprising peptide or polypeptide to facilitate purification. For example, an expression vector can include an epitope-encoding nucleic acid sequence linked to six histidine residues followed by a thioredoxin and an enterokinase cleavage site (see e.g., Williams (1995) Biochemistry 34:1787-1797; Dobeli (1998) Protein Expr. Purif. 12:404-414). The histidine residues facilitate detection and purification while the enterokinase cleavage site provides a means for purifying the epitope from the remainder of the fusion protein. Technology pertaining to vectors encoding fusion proteins and application of fusion proteins are well described in the scientific and patent literature, see e.g., Kroll (1993) DNA Cell. Biol., 12:441-53.

Transcriptional and translational control sequences

The invention provides nucleic acid (e.g., DNA) sequences of the invention operatively linked to expression (e.g., transcriptional or translational) control sequence(s), e.g., promoters or enhancers, to direct or modulate RNA synthesis/ expression. The expression control sequence can be in an expression vector. Exemplary bacterial promoters include lacI, lacZ, T3, T7, gpt, lambda PR, PL and trp. Exemplary eukaryotic promoters include CMV immediate early, HSV thymidine kinase, early and late SV40, LTRs from retrovirus, and mouse metallothionein I.

Promoters suitable for expressing a polypeptide in bacteria include the *E. coli* lac or trp promoters, the lacI promoter, the lacZ promoter, the T3 promoter, the T7 promoter, the gpt promoter, the lambda PR promoter, the lambda PL promoter, promoters from operons encoding glycolytic enzymes such as 3-phosphoglycerate kinase (PGK), and the acid phosphatase promoter. Eukaryotic promoters include the CMV immediate early promoter, the HSV thymidine kinase promoter, heat shock promoters, the early and late SV40 promoter, LTRs from retroviruses, and the mouse metallothionein-I promoter. Other promoters known to control expression of genes in prokaryotic or eukaryotic cells or their viruses may also be used.

Tissue-Specific Plant Promoters

The invention provides expression cassettes that can be expressed in a tissue-specific manner, e.g., that can express an amylase of the invention in a tissue-

specific manner. The invention also provides plants or seeds that express an amylase of the invention in a tissue-specific manner. The tissue-specificity can be seed specific, stem specific, leaf specific, root specific, fruit specific and the like.

In one aspect, a constitutive promoter such as the CaMV 35S promoter can be used for expression in specific parts of the plant or seed or throughout the plant. For example, for overexpression, a plant promoter fragment can be employed which will direct expression of a nucleic acid in some or all tissues of a plant, e.g., a regenerated plant. Such promoters are referred to herein as "constitutive" promoters and are active under most environmental conditions and states of development or cell differentiation. Examples of constitutive promoters include the cauliflower mosaic virus (CaMV) 35S transcription initiation region, the 1'- or 2'- promoter derived from T-DNA of *Agrobacterium tumefaciens*, and other transcription initiation regions from various plant genes known to those of skill. Such genes include, e.g., *ACT11* from *Arabidopsis* (Huang (1996) *Plant Mol. Biol.* 33:125-139); *Cat3* from *Arabidopsis* (GenBank No. U43147, Zhong (1996) *Mol. Gen. Genet.* 251:196-203); the gene encoding stearoyl-acyl carrier protein desaturase from *Brassica napus* (Genbank No. X74782, Solocombe (1994) *Plant Physiol.* 104:1167-1176); *GPc1* from maize (GenBank No. X15596; Martinez (1989) *J. Mol. Biol* 208:551-565); the *Gpc2* from maize (GenBank No. U45855, Manjunath (1997) *Plant Mol. Biol.* 33:97-112); plant promoters described in U.S. Patent Nos. 4,962,028; 5,633,440.

The invention uses tissue-specific or constitutive promoters derived from viruses which can include, *e.g.*, the tobamovirus subgenomic promoter (Kumagai (1995) Proc. Natl. Acad. Sci. USA 92:1679-1683; the rice tungro bacilliform virus (RTBV), which replicates only in phloem cells in infected rice plants, with its promoter which drives strong phloem-specific reporter gene expression; the cassava vein mosaic virus (CVMV) promoter, with highest activity in vascular elements, in leaf mesophyll cells, and in root tips (Verdaguer (1996) Plant Mol. Biol. 31:1129-1139).

Alternatively, the plant promoter may direct expression of amylase-expressing nucleic acid in a specific tissue, organ or cell type (*i.e.* tissue-specific promoters) or may be otherwise under more precise environmental or developmental control or under the control of an inducible promoter. Examples of environmental conditions that may affect transcription include anaerobic conditions, elevated temperature, the presence of light, or sprayed with chemicals/hormones. For example, the

invention incorporates the drought-inducible promoter of maize (Busk (1997) supra); the cold, drought, and high salt inducible promoter from potato (Kirch (1997) Plant Mol. Biol. 33:897 909).

Tissue-specific promoters can promote transcription only within a certain time frame of developmental stage within that tissue. See, e.g., Blazquez (1998) Plant Cell 10:791-800, characterizing the *Arabidopsis* LEAFY gene promoter. See also Cardon (1997) *Plant J* 12:367-77, describing the transcription factor SPL3, which recognizes a conserved sequence motif in the promoter region of the *A. thaliana* floral meristem identity gene AP1; and Mandel (1995) Plant Molecular Biology, Vol. 29, pp 995-1004, describing the meristem promoter eIF4. Tissue specific promoters which are active throughout the life cycle of a particular tissue can be used. In one aspect, the nucleic acids of the invention are operably linked to a promoter active primarily only in cotton fiber cells. In one aspect, the nucleic acids of the invention are operably linked to a promoter active primarily during the stages of cotton fiber cell elongation, e.g., as described by Rinehart (1996) supra. The nucleic acids can be operably linked to the Fbl2A gene promoter to be preferentially expressed in cotton fiber cells (Ibid) . See also, John (1997) Proc. Natl. Acad. Sci. USA 89:5769-5773; John, et al., U.S. Patent Nos. 5,608,148 and 5,602,321, describing cotton fiber-specific promoters and methods for the construction of transgenic cotton plants. Root-specific promoters may also be used to express the nucleic acids of the invention. Examples of root-specific promoters include the promoter from the alcohol dehydrogenase gene (DeLisle (1990) Int. Rev. Cytol. 123:39-60). Other promoters that can be used to express the nucleic acids of the invention include, e.g., ovule-specific, embryo-specific, endosperm-specific, integument-specific, seed coat-specific promoters, or some combination thereof; a leaf-specific promoter (see, e.g., Busk (1997) Plant J. 11:1285 1295, describing a leaf-specific promoter in maize); the ORF13 promoter from *Agrobacterium rhizogenes* (which exhibits high activity in roots, see, e.g., Hansen (1997) supra); a maize pollen specific promoter (see, e.g., Guerrero (1990) Mol. Gen. Genet. 224:161 168); a tomato promoter active during fruit ripening, senescence and abscission of leaves and, to a lesser extent, of flowers can be used (see, e.g., Blume (1997) Plant J. 12:731 746); a pistil-specific promoter from the potato SK2 gene (see, e.g., Ficker (1997) Plant Mol. Biol. 35:425 431); the Blec4 gene from pea, which is active in epidermal tissue of vegetative and floral shoot apices of transgenic alfalfa making it a useful tool to target the expression of

foreign genes to the epidermal layer of actively growing shoots or fibers; the ovule-specific BEL1 gene (see, e.g., Reiser (1995) Cell 83:735-742, GenBank No. U39944); and/or, the promoter in Klee, U.S. Patent No. 5,589,583, describing a plant promoter region is capable of conferring high levels of transcription in meristematic tissue and/or rapidly dividing cells.

Alternatively, plant promoters which are inducible upon exposure to plant hormones, such as auxins, are used to express the nucleic acids of the invention. For example, the invention can use the auxin-response elements E1 promoter fragment (AuxREs) in the soybean (*Glycine max* L.) (Liu (1997) Plant Physiol. 115:397-407); the auxin-responsive *Arabidopsis* GST6 promoter (also responsive to salicylic acid and hydrogen peroxide) (Chen (1996) Plant J. 10: 955-966); the auxin-inducible parC promoter from tobacco (Sakai (1996) 37:906-913); a plant biotin response element (Streit (1997) Mol. Plant Microbe Interact. 10:933-937); and, the promoter responsive to the stress hormone abscisic acid (Sheen (1996) Science 274:1900-1902).

The nucleic acids of the invention can also be operably linked to plant promoters which are inducible upon exposure to chemicals reagents which can be applied to the plant, such as herbicides or antibiotics. For example, the maize In2-2 promoter, activated by benzenesulfonamide herbicide safeners, can be used (De Veylder (1997) Plant Cell Physiol. 38:568-577); application of different herbicide safeners induces distinct gene expression patterns, including expression in the root, hydathodes, and the shoot apical meristem. Coding sequence can be under the control of, *e.g.*, a tetracycline-inducible promoter, *e.g.*, as described with transgenic tobacco plants containing the *Avena sativa* L. (oat) arginine decarboxylase gene (Masgrau (1997) Plant J. 11:465-473); or, a salicylic acid-responsive element (Stange (1997) Plant J. 11:1315-1324). Using chemically- (*e.g.*, hormone- or pesticide-) induced promoters, *i.e.*, promoter responsive to a chemical which can be applied to the transgenic plant in the field, expression of a polypeptide of the invention can be induced at a particular stage of development of the plant. Thus, the invention also provides for transgenic plants containing an inducible gene encoding for polypeptides of the invention whose host range is limited to target plant species, such as corn, rice, barley, wheat, potato or other crops, inducible at any stage of development of the crop.

One of skill will recognize that a tissue-specific plant promoter may drive expression of operably linked sequences in tissues other than the target tissue. Thus, a

tissue-specific promoter is one that drives expression preferentially in the target tissue or cell type, but may also lead to some expression in other tissues as well.

The nucleic acids of the invention can also be operably linked to plant promoters which are inducible upon exposure to chemicals reagents. These reagents include, e.g., herbicides, synthetic auxins, or antibiotics which can be applied, e.g., sprayed, onto transgenic plants. Inducible expression of the amylase-producing nucleic acids of the invention will allow the grower to select plants with the optimal starch / sugar ratio. The development of plant parts can thus controlled. In this way the invention provides the means to facilitate the harvesting of plants and plant parts. For example, in various embodiments, the maize In2-2 promoter, activated by benzenesulfonamide herbicide safeners, is used (De Veylder (1997) Plant Cell Physiol. 38:568-577); application of different herbicide safeners induces distinct gene expression patterns, including expression in the root, hydathodes, and the shoot apical meristem. Coding sequences of the invention are also under the control of a tetracycline-inducible promoter, e.g., as described with transgenic tobacco plants containing the *Avena sativa* L. (oat) arginine decarboxylase gene (Masgrau (1997) Plant J. 11:465-473); or, a salicylic acid-responsive element (Stange (1997) Plant J. 11:1315-1324).

If proper polypeptide expression is desired, a polyadenylation region at the 3'-end of the coding region should be included. The polyadenylation region can be derived from the natural gene, from a variety of other plant genes, or from genes in the *Agrobacterial* T-DNA.

Expression vectors and cloning vehicles

The invention provides expression vectors and cloning vehicles comprising nucleic acids of the invention, e.g., sequences encoding the amylases of the invention. Expression vectors and cloning vehicles of the invention can comprise viral particles, baculovirus, phage, plasmids, phagemids, cosmids, fosmids, bacterial artificial chromosomes, viral DNA (e.g., vaccinia, adenovirus, foul pox virus, pseudorabies and derivatives of SV40), P1-based artificial chromosomes, yeast plasmids, yeast artificial chromosomes, and any other vectors specific for specific hosts of interest (such as bacillus, Aspergillus and yeast). Vectors of the invention can include chromosomal, non-chromosomal and synthetic DNA sequences. Large numbers of suitable vectors are known to those of skill in the art, and are commercially available. Exemplary vectors are

include: bacterial: pQE vectors (Qiagen), pBluescript plasmids, pNH vectors, (lambda-ZAP vectors (Stratagene); ptrc99a, pKK223-3, pDR540, pRIT2T (Pharmacia); Eukaryotic: pXT1, pSG5 (Stratagene), pSVK3, pBPV, pMSG, pSVLSV40 (Pharmacia). However, any other plasmid or other vector may be used so long as they are replicable and viable in the host. Low copy number or high copy number vectors may be employed with the present invention.

The expression vector can comprise a promoter, a ribosome binding site for translation initiation and a transcription terminator. The vector may also include appropriate sequences for amplifying expression. Mammalian expression vectors can comprise an origin of replication, any necessary ribosome binding sites, a polyadenylation site, splice donor and acceptor sites, transcriptional termination sequences, and 5' flanking non-transcribed sequences. In some aspects, DNA sequences derived from the SV40 splice and polyadenylation sites may be used to provide the required non-transcribed genetic elements.

In one aspect, the expression vectors contain one or more selectable marker genes to permit selection of host cells containing the vector. Such selectable markers include genes encoding dihydrofolate reductase or genes conferring neomycin resistance for eukaryotic cell culture, genes conferring tetracycline or ampicillin resistance in *E. coli*, and the *S. cerevisiae* TRP1 gene. Promoter regions can be selected from any desired gene using chloramphenicol transferase (CAT) vectors or other vectors with selectable markers.

Vectors for expressing the polypeptide or fragment thereof in eukaryotic cells can also contain enhancers to increase expression levels. Enhancers are cis-acting elements of DNA, usually from about 10 to about 300 bp in length that act on a promoter to increase its transcription. Examples include the SV40 enhancer on the late side of the replication origin bp 100 to 270, the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and the adenovirus enhancers.

A nucleic acid sequence can be inserted into a vector by a variety of procedures. In general, the sequence is ligated to the desired position in the vector following digestion of the insert and the vector with appropriate restriction endonucleases. Alternatively, blunt ends in both the insert and the vector may be ligated. A variety of cloning techniques are known in the art, e.g., as described in Ausubel and

Sambrook. Such procedures and others are deemed to be within the scope of those skilled in the art.

The vector can be in the form of a plasmid, a viral particle, or a phage. Other vectors include chromosomal, non-chromosomal and synthetic DNA sequences, derivatives of SV40; bacterial plasmids, phage DNA, baculovirus, yeast plasmids, vectors derived from combinations of plasmids and phage DNA, viral DNA such as vaccinia, adenovirus, fowl pox virus, and pseudorabies. A variety of cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described by, e.g., Sambrook.

Particular bacterial vectors which can be used include the commercially available plasmids comprising genetic elements of the well known cloning vector pBR322 (ATCC 37017), pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Sweden), GEM1 (Promega Biotec, Madison, WI, USA) pQE70, pQE60, pQE-9 (Qiagen), pD10, psiX174 pBluescript II KS, pNH8A, pNH16a, pNH18A, pNH46A (Stratagene), ptrc99a, pKK223-3, pKK233-3, DR540, pRIT5 (Pharmacia), pKK232-8 and pCM7. Particular eukaryotic vectors include pSV2CAT, pOG44, pXT1, pSG (Stratagene) pSVK3, pBPV, pMSG, and pSVL (Pharmacia). However, any other vector may be used as long as it is replicable and viable in the host cell.

The nucleic acids of the invention can be expressed in expression cassettes, vectors or viruses and transiently or stably expressed in plant cells and seeds. One exemplary transient expression system uses episomal expression systems, e.g., cauliflower mosaic virus (CaMV) viral RNA generated in the nucleus by transcription of an episomal mini-chromosome containing supercoiled DNA, see, e.g., Covey (1990) Proc. Natl. Acad. Sci. USA 87:1633-1637. Alternatively, coding sequences, i.e., all or sub-fragments of sequences of the invention can be inserted into a plant host cell genome becoming an integral part of the host chromosomal DNA. Sense or antisense transcripts can be expressed in this manner. A vector comprising the sequences (e.g., promoters or coding regions) from nucleic acids of the invention can comprise a marker gene that confers a selectable phenotype on a plant cell or a seed. For example, the marker may encode biocide resistance, particularly antibiotic resistance, such as resistance to kanamycin, G418, bleomycin, hygromycin, or herbicide resistance, such as resistance to chlorosulfuron or Basta.

Expression vectors capable of expressing nucleic acids and proteins in plants are well known in the art, and can include, *e.g.*, vectors from *Agrobacterium* spp.,

potato virus X (see, *e.g.*, Angell (1997) EMBO J. 16:3675-3684), tobacco mosaic virus (see, *e.g.*, Casper (1996) Gene 173:69-73), tomato bushy stunt virus (see, *e.g.*, Hillman (1989) Virology 169:42-50), tobacco etch virus (see, *e.g.*, Dolja (1997) Virology 234:243-252), bean golden mosaic virus (see, *e.g.*, Morinaga (1993) Microbiol Immunol. 37:471-476), cauliflower mosaic virus (see, *e.g.*, Cecchini (1997) Mol. Plant Microbe Interact. 10:1094-1101), maize Ac/Ds transposable element (see, *e.g.*, Rubin (1997) Mol. Cell. Biol. 17:6294-6302; Kunze (1996) Curr. Top. Microbiol. Immunol. 204:161-194), and the maize suppressor-mutator (Spm) transposable element (see, *e.g.*, Schlappi (1996) Plant Mol. Biol. 32:717-725); and derivatives thereof.

In one aspect, the expression vector can have two replication systems to allow it to be maintained in two organisms, for example in mammalian or insect cells for expression and in a prokaryotic host for cloning and amplification. Furthermore, for integrating expression vectors, the expression vector can contain at least one sequence homologous to the host cell genome. It can contain two homologous sequences which flank the expression construct. The integrating vector can be directed to a specific locus in the host cell by selecting the appropriate homologous sequence for inclusion in the vector. Constructs for integrating vectors are well known in the art.

Expression vectors of the invention may also include a selectable marker gene to allow for the selection of bacterial strains that have been transformed, e.g., genes which render the bacteria resistant to drugs such as ampicillin, chloramphenicol, erythromycin, kanamycin, neomycin and tetracycline. Selectable markers can also include biosynthetic genes, such as those in the histidine, tryptophan and leucine biosynthetic pathways.

Host cells and transformed cells

The invention also provides a transformed cell comprising a nucleic acid sequence of the invention, e.g., a sequence encoding an amylase of the invention, or a vector of the invention. The host cell may be any of the host cells familiar to those skilled in the art, including prokaryotic cells, eukaryotic cells, such as bacterial cells, fungal cells, yeast cells, mammalian cells, insect cells, or plant cells. Exemplary bacterial cells include *E. coli*, any *Streptomyces* or *Bacillus* (e.g., *Bacillus cereus, Bacillus subtilis*), *Salmonella typhimurium* and various species within the genera *Bacillus*, *Streptomyces*, and *Staphylococcus*. Exemplary insect cells include *Drosophila S2* and

Spodoptera Sf9. Exemplary animal cells include CHO, COS or Bowes melanoma or any mouse or human cell line. The selection of an appropriate host is within the abilities of those skilled in the art. Techniques for transforming a wide variety of higher plant species are well known and described in the technical and scientific literature. See, e.g., Weising (1988) Ann. Rev. Genet. 22:421-477, U.S. Patent No. 5,750,870.

The vector can be introduced into the host cells using any of a variety of techniques, including transformation, transfection, transduction, viral infection, gene guns, or Ti-mediated gene transfer. Particular methods include calcium phosphate transfection, DEAE-Dextran mediated transfection, lipofection, or electroporation (Davis, L., Dibner, M., Battey, I., Basic Methods in Molecular Biology, (1986)).

In one aspect, the nucleic acids or vectors of the invention are introduced into the cells for screening, thus, the nucleic acids enter the cells in a manner suitable for subsequent expression of the nucleic acid. The method of introduction is largely dictated by the targeted cell type. Exemplary methods include $CaPO_4$ precipitation, liposome fusion, lipofection (e.g., LIPOFECTIN™), electroporation, viral infection, etc. The candidate nucleic acids may stably integrate into the genome of the host cell (for example, with retroviral introduction) or may exist either transiently or stably in the cytoplasm (i.e. through the use of traditional plasmids, utilizing standard regulatory sequences, selection markers, etc.). As many pharmaceutically important screens require human or model mammalian cell targets, retroviral vectors capable of transfecting such targets are preferred.

Where appropriate, the engineered host cells can be cultured in conventional nutrient media modified as appropriate for activating promoters, selecting transformants or amplifying the genes of the invention. Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, the selected promoter may be induced by appropriate means (e.g., temperature shift or chemical induction) and the cells may be cultured for an additional period to allow them to produce the desired polypeptide or fragment thereof.

Cells can be harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract is retained for further purification. Microbial cells employed for expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents. Such methods are well known to those skilled in the art. The expressed

polypeptide or fragment thereof can be recovered and purified from recombinant cell cultures by methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Protein refolding steps can be used, as necessary, in completing configuration of the polypeptide. If desired, high performance liquid chromatography (HPLC) can be employed for final purification steps.

Various mammalian cell culture systems can also be employed to express recombinant protein. Examples of mammalian expression systems include the COS-7 lines of monkey kidney fibroblasts and other cell lines capable of expressing proteins from a compatible vector, such as the C127, 3T3, CHO, HeLa and BHK cell lines.

The constructs in host cells can be used in a conventional manner to produce the gene product encoded by the recombinant sequence. Depending upon the host employed in a recombinant production procedure, the polypeptides produced by host cells containing the vector may be glycosylated or may be non-glycosylated. Polypeptides of the invention may or may not also include an initial methionine amino acid residue.

Cell-free translation systems can also be employed to produce a polypeptide of the invention. Cell-free translation systems can use mRNAs transcribed from a DNA construct comprising a promoter operably linked to a nucleic acid encoding the polypeptide or fragment thereof. In some aspects, the DNA construct may be linearized prior to conducting an in vitro transcription reaction. The transcribed mRNA is then incubated with an appropriate cell-free translation extract, such as a rabbit reticulocyte extract, to produce the desired polypeptide or fragment thereof.

The expression vectors can contain one or more selectable marker genes to provide a phenotypic trait for selection of transformed host cells such as dihydrofolate reductase or neomycin resistance for eukaryotic cell culture, or such as tetracycline or ampicillin resistance in *E. coli*.

Amplification of Nucleic Acids

In practicing the invention, nucleic acids of the invention and nucleic acids encoding the polypeptides of the invention, or modified nucleic acids of the invention, can be reproduced by amplification. Amplification can also be used to clone or modify

the nucleic acids of the invention. Thus, the invention provides amplification primer sequence pairs for amplifying nucleic acids of the invention. One of skill in the art can design amplification primer sequence pairs for any part of or the full length of these sequences.

Amplification reactions can also be used to quantify the amount of nucleic acid in a sample (such as the amount of message in a cell sample), label the nucleic acid (e.g., to apply it to an array or a blot), detect the nucleic acid, or quantify the amount of a specific nucleic acid in a sample. In one aspect of the invention, message isolated from a cell or a cDNA library are amplified.

The skilled artisan can select and design suitable oligonucleotide amplification primers. Amplification methods are also well known in the art, and include, e.g., polymerase chain reaction, PCR (see, e.g., PCR PROTOCOLS, A GUIDE TO METHODS AND APPLICATIONS, ed. Innis, Academic Press, N.Y. (1990) and PCR STRATEGIES (1995), ed. Innis, Academic Press, Inc., N.Y., ligase chain reaction (LCR) (see, e.g., Wu (1989) Genomics 4:560; Landegren (1988) Science 241:1077; Barringer (1990) Gene 89:117); transcription amplification (see, e.g., Kwoh (1989) Proc. Natl. Acad. Sci. USA 86:1173); and, self-sustained sequence replication (see, e.g., Guatelli (1990) Proc. Natl. Acad. Sci. USA 87:1874); Q Beta replicase amplification (see, e.g., Smith (1997) J. Clin. Microbiol. 35:1477-1491), automated Q-beta replicase amplification assay (see, e.g., Burg (1996) Mol. Cell. Probes 10:257-271) and other RNA polymerase mediated techniques (e.g., NASBA, Cangene, Mississauga, Ontario); see also Berger (1987) Methods Enzymol. 152:307-316; Sambrook; Ausubel; U.S. Patent Nos. 4,683,195 and 4,683,202; Sooknanan (1995) Biotechnology 13:563-564.

Determining the degree of sequence identity

The invention provides nucleic acids comprising sequences having at least about 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more, or complete (100%) sequence identity to an exemplary nucleic acid of the invention over a region of at least about 50, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400, 1450, 1500, 1550 or more, residues.

The invention provides polypeptides comprising sequences having at least about 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more, or complete (100%) sequence identity to an exemplary polypeptide of the invention. The extent of sequence identity (homology) may be determined using any computer program and associated parameters, including those described herein, such as BLAST 2.2.2. or FASTA version 3.0t78, with the default parameters.

Figure 35 is a chart describing selected characteristics of exemplary nucleic acids and polypeptides of the invention, including sequence identity comparison of the exemplary sequences to public databases. All sequences described in Figure 35 have been subject to a BLAST search (as described in detail, below) against two sets of databases. The first database set is available through NCBI (National Center for Biotechnology Information). All results from searches against these databases are found in the columns entitled "NR Description", "NR Accession Code", "NR Evalue" or "NR Organism". "NR" refers to the Non-Redundant nucleotide database maintained by NCBI. This database is a composite of GenBank, GenBank updates, and EMBL updates. The entries in the column "NR Description" refer to the definition line in any given NCBI record, which includes a description of the sequence, such as the source organism, gene name/protein name, or some description of the function of the sequence. The entries in the column "NR Accession Code" refer to the unique identifier given to a sequence record. The entries in the column "NR Evalue" refer to the Expect value (Evalue), which represents the probability that an alignment score as good as the one found between the query sequence (the sequences of the invention) and a database sequence would be found in the same number of comparisons between random sequences as was done in the present BLAST search. The entries in the column "NR Organism" refer to the source organism of the sequence identified as the closest BLAST hit. The second set of databases is collectively known as the Geneseq™ database, which is available through Thomson Derwent (Philadelphia, PA). All results from searches against this database are found in the columns entitled "Geneseq Protein Description", "Geneseq Protein Accession Code", "Geneseq Protein Evalue", "Geneseq DNA Description", "Geneseq DNA Accession Code" or "Geneseq DNA Evalue". The information found in these

columns is comparable to the information found in the NR columns described above, except that it was derived from BLAST searches against the Geneseq™ database instead of the NCBI databases. In addition, this table includes the column "Predicted EC No.". An EC number is the number assigned to a type of enzyme according to a scheme of standardized enzyme nomenclature developed by the Enzyme Commission of the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology (IUBMB). The results in the "Predicted EC No." column are determined by a BLAST search against the Kegg (Kyoto Encyclopedia of Genes and Genomes) database. If the top BLAST match has an Evalue equal to or less than e^{-6}, the EC number assigned to the top match is entered into the table. The EC number of the top hit is used as a guide to what the EC number of the sequence of the invention might be. The columns "Query DNA Length" and "Query Protein Length" refer to the number of nucleotides or the number amino acids, respectively, in the sequence of the invention that was searched or queried against either the NCBI or Geneseq databases. The columns "Geneseq or NR DNA Length" and "Geneseq or NR Protein Length" refer to the number of nucleotides or the number amino acids, respectively, in the sequence of the top match from the BLAST search. The results provided in these columns are from the search that returned the lower Evalue, either from the NCBI databases or the Geneseq database. The columns "Geneseq or NR %ID Protein" and "Geneseq or NR %ID DNA" refer to the percent sequence identity between the sequence of the invention and the sequence of the top BLAST match. The results provided in these columns are from the search that returned the lower Evalue, either from the NCBI databases or the Geneseq database.

Homologous sequences also include RNA sequences in which uridines replace the thymines in the nucleic acid sequences. The homologous sequences may be obtained using any of the procedures described herein or may result from the correction of a sequencing error. It will be appreciated that the nucleic acid sequences as set forth herein can be represented in the traditional single character format (see, e.g., Stryer, Lubert. Biochemistry, 3rd Ed., W. H Freeman & Co., New York) or in any other format which records the identity of the nucleotides in a sequence.

Various sequence comparison programs identified herein are used in this aspect of the invention. Protein and/or nucleic acid sequence identities (homologies) may be evaluated using any of the variety of sequence comparison algorithms and programs known in the art. Such algorithms and programs include, but are not limited to,

TBLASTN, BLASTP, FASTA, TFASTA, and CLUSTALW (Pearson and Lipman, Proc. Natl. Acad. Sci. USA 85(8):2444-2448, 1988; Altschul et al., J. Mol. Biol. 215(3):403-410, 1990; Thompson et al., Nucleic Acids Res. 22(2):4673-4680, 1994; Higgins et al., Methods Enzymol. 266:383-402, 1996; Altschul et al., J. Mol. Biol. 215(3):403-410, 1990; Altschul et al., Nature Genetics 3:266-272, 1993).

Homology or sequence identity can be measured using sequence analysis software (e.g., Sequence Analysis Software Package of the Genetics Computer Group, University of Wisconsin Biotechnology Center, 1710 University Avenue, Madison, WI 53705). Such software matches similar sequences by assigning degrees of homology to various deletions, substitutions and other modifications. The terms "homology" and "identity" in the context of two or more nucleic acids or polypeptide sequences, refer to two or more sequences or subsequences that are the same or have a specified percentage of amino acid residues or nucleotides that are the same when compared and aligned for maximum correspondence over a comparison window or designated region as measured using any number of sequence comparison algorithms or by manual alignment and visual inspection. For sequence comparison, one sequence can act as a reference sequence, e.g., a sequence of the invention, to which test sequences are compared. When using a sequence comparison algorithm, test and reference sequences are entered into a computer, subsequence coordinates are designated, if necessary, and sequence algorithm program parameters are designated. Default program parameters can be used, or alternative parameters can be designated. The sequence comparison algorithm then calculates the percent sequence identities for the test sequences relative to the reference sequence, based on the program parameters.

A "comparison window", as used herein, includes reference to a segment of any one of the numbers of contiguous residues. For example, in alternative aspects of the invention, contiguous residues ranging anywhere from 20 to the full length of an exemplary polypeptide or nucleic acid sequence of the invention are compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned. If the reference sequence has the requisite sequence identity to an exemplary polypeptide or nucleic acid sequence of the invention, e.g., 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%,

98%, 99%, or more sequence identity to a sequence of the invention, that sequence is within the scope of the invention. In alternative embodiments, subsequences ranging from about 20 to 600, about 50 to 200, and about 100 to 150 are compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned.

Methods of alignment of sequence for comparison are well known in the art. In alternative aspects, optimal alignment of sequences for comparison can be conducted, e.g., by the local homology algorithm of Smith & Waterman, Adv. Appl. Math. 2:482, 1981, by the homology alignment algorithm of Needleman & Wunsch, J. Mol. Biol. 48:443, 1970, by the search for similarity method of person & Lipman, Proc. Nat'l. Acad. Sci. USA 85:2444, 1988, by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), or by manual alignment and visual inspection. Other algorithms for determining homology or identity include, for example, in addition to a BLAST program (Basic Local Alignment Search Tool at the National Center for Biological Information), ALIGN, AMAS (Analysis of Multiply Aligned Sequences), AMPS (Protein Multiple Sequence Alignment), ASSET (Aligned Segment Statistical Evaluation Tool), BANDS, BESTSCOR, BIOSCAN (Biological Sequence Comparative Analysis Node), BLIMPS (BLocks IMProved Searcher), FASTA, Intervals & Points, BMB, CLUSTAL V, CLUSTAL W, CONSENSUS, LCONSENSUS, WCONSENSUS, Smith-Waterman algorithm, DARWIN, Las Vegas algorithm, FNAT (Forced Nucleotide Alignment Tool), Framealign, Framesearch, DYNAMIC, FILTER, FSAP (Fristensky Sequence Analysis Package), GAP (Global Alignment Program), GENAL, GIBBS, GenQuest, ISSC (Sensitive Sequence Comparison), LALIGN (Local Sequence Alignment), LCP (Local Content Program), MACAW (Multiple Alignment Construction & Analysis Workbench), MAP (Multiple Alignment Program), MBLKP, MBLKN, PIMA (Pattern-Induced Multi-sequence Alignment), SAGA (Sequence Alignment by Genetic Algorithm) and WHAT-IF. Such alignment programs can also be used to screen genome databases to identify polynucleotide sequences having substantially identical sequences. A number of genome databases are available, for example, a substantial portion of the human genome is available as part of the Human Genome Sequencing Project (Gibbs, 1995). Several genomes have been sequenced, e.g., *M. genitalium* (Fraser et al., 1995), *M. jannaschii*

(Bult et al., 1996), *H. influenzae* (Fleischmann et al., 1995), *E. coli* (Blattner et al., 1997), and yeast (*S. cerevisiae*) (Mewes et al., 1997), and *D. melanogaster* (Adams et al., 2000). Significant progress has also been made in sequencing the genomes of model organism, such as mouse, *C. elegans*, and *Arabadopsis* sp. Databases containing genomic information annotated with some functional information are maintained by different organization, and are accessible via the internet.

BLAST, BLAST 2.0 and BLAST 2.2.2 algorithms also can be used to practice the invention. They are described, e.g., in Altschul (1977) Nuc. Acids Res. 25:3389-3402; Altschul (1990) J. Mol. Biol. 215:403-410. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information. This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul (1990) supra). These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always >0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) of 10, M=5, N=-4 and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength of 3, and expectations (E) of 10, and the BLOSUM62 scoring matrix (see Henikoff & Henikoff (1989) Proc. Natl. Acad. Sci. USA 89:10915) alignments (B) of 50, expectation (E) of 10, M=5, N= -4, and a comparison of both strands. The BLAST algorithm also performs a statistical analysis of the similarity between two sequences (see, e.g., Karlin & Altschul (1993) Proc. Natl. Acad. Sci. USA 90:5873). One measure of similarity provided by BLAST algorithm is the smallest sum

probability (P(N)), which provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance. For example, a nucleic acid is considered similar to a references sequence if the smallest sum probability in a comparison of the test nucleic acid to the reference nucleic acid is less than about 0.2, more preferably less than about 0.01, and most preferably less than about 0.001. In one aspect, protein and nucleic acid sequence homologies are evaluated using the Basic Local Alignment Search Tool ("BLAST"). For example, five specific BLAST programs can be used to perform the following task: (1) BLASTP and BLAST3 compare an amino acid query sequence against a protein sequence database; (2) BLASTN compares a nucleotide query sequence against a nucleotide sequence database; (3) BLASTX compares the six-frame conceptual translation products of a query nucleotide sequence (both strands) against a protein sequence database; (4) TBLASTN compares a query protein sequence against a nucleotide sequence database translated in all six reading frames (both strands); and, (5) TBLASTX compares the six-frame translations of a nucleotide query sequence against the six-frame translations of a nucleotide sequence database. The BLAST programs identify homologous sequences by identifying similar segments, which are referred to herein as "high-scoring segment pairs," between a query amino or nucleic acid sequence and a test sequence which is preferably obtained from a protein or nucleic acid sequence database. High-scoring segment pairs are preferably identified (i.e., aligned) by means of a scoring matrix, many of which are known in the art. Preferably, the scoring matrix used is the BLOSUM62 matrix (Gonnet et al., Science 256:1443-1445, 1992; Henikoff and Henikoff, Proteins 17:49-61, 1993). Less preferably, the PAM or PAM250 matrices may also be used (see, e.g., Schwartz and Dayhoff, eds., 1978, Matrices for Detecting Distance Relationships: Atlas of Protein Sequence and Structure, Washington: National Biomedical Research Foundation).

In one aspect of the invention, to determine if a nucleic acid has the requisite sequence identity to be within the scope of the invention, the NCBI BLAST 2.2.2 programs is used, default options to blastp. There are about 38 setting options in the BLAST 2.2.2 program. In this exemplary aspect of the invention, all default values are used except for the default filtering setting (i.e., all parameters set to default except filtering which is set to OFF); in its place a "-F F" setting is used, which disables filtering. Use of default filtering often results in Karlin-Altschul violations due to short length of sequence.

The default values used in this exemplary aspect of the invention, and to determine the values in Figure 35, as discussed above, include:

"Filter for low complexity: ON

Word Size: 3

Matrix: Blosum62

Gap Costs: Existence:11

Extension:1"

Other default settings can be: filter for low complexity OFF, word size of 3 for protein, BLOSUM62 matrix, gap existence penalty of -11 and a gap extension penalty of -1. An exemplary NCBI BLAST 2.2.2 program setting has the "-W" option default to 0. This means that, if not set, the word size defaults to 3 for proteins and 11 for nucleotides.

Computer systems and computer program products

To determine and identify sequence identities, structural homologies, motifs and the like in silico, the sequence of the invention can be stored, recorded, and manipulated on any medium which can be read and accessed by a computer. Accordingly, the invention provides computers, computer systems, computer readable mediums, computer programs products and the like recorded or stored thereon the nucleic acid and polypeptide sequences of the invention. As used herein, the words "recorded" and "stored" refer to a process for storing information on a computer medium. A skilled artisan can readily adopt any known methods for recording information on a computer readable medium to generate manufactures comprising one or more of the nucleic acid and/or polypeptide sequences of the invention.

Another aspect of the invention is a computer readable medium having recorded thereon at least one nucleic acid and/or polypeptide sequence of the invention. Computer readable media include magnetically readable media, optically readable media, electronically readable media and magnetic/optical media. For example, the computer readable media may be a hard disk, a floppy disk, a magnetic tape, CD-ROM, Digital Versatile Disk (DVD), Random Access Memory (RAM), or Read Only Memory (ROM) as well as other types of other media known to those skilled in the art.

Aspects of the invention include systems (e.g., internet based systems), particularly computer systems, which store and manipulate the sequences and sequence

information described herein. One example of a computer system 100 is illustrated in block diagram form in Figure 1. As used herein, "a computer system" refers to the hardware components, software components, and data storage components used to analyze a nucleotide or polypeptide sequence of the invention. The computer system 100 can include a processor for processing, accessing and manipulating the sequence data. The processor 105 can be any well-known type of central processing unit, such as, for example, the Pentium III from Intel Corporation, or similar processor from Sun, Motorola, Compaq, AMD or International Business Machines. The computer system 100 is a general purpose system that comprises the processor 105 and one or more internal data storage components 110 for storing data, and one or more data retrieving devices for retrieving the data stored on the data storage components. A skilled artisan can readily appreciate that any one of the currently available computer systems are suitable.

In one aspect, the computer system 100 includes a processor 105 connected to a bus which is connected to a main memory 115 (preferably implemented as RAM) and one or more internal data storage devices 110, such as a hard drive and/or other computer readable media having data recorded thereon. The computer system 100 can further include one or more data retrieving device 118 for reading the data stored on the internal data storage devices 110. The data retrieving device 118 may represent, for example, a floppy disk drive, a compact disk drive, a magnetic tape drive, or a modem capable of connection to a remote data storage system (e.g., via the internet) etc. In some embodiments, the internal data storage device 110 is a removable computer readable medium such as a floppy disk, a compact disk, a magnetic tape, etc. containing control logic and/or data recorded thereon. The computer system 100 may advantageously include or be programmed by appropriate software for reading the control logic and/or the data from the data storage component once inserted in the data retrieving device. The computer system 100 includes a display 120 which is used to display output to a computer user. It should also be noted that the computer system 100 can be linked to other computer systems 125a-c in a network or wide area network to provide centralized access to the computer system 100. Software for accessing and processing the nucleotide or amino acid sequences of the invention can reside in main memory 115 during execution. In some aspects, the computer system 100 may further comprise a sequence comparison algorithm for comparing a nucleic acid sequence of the invention. The algorithm and sequence(s) can be stored on a computer readable medium. A "sequence

comparison algorithm" refers to one or more programs which are implemented (locally or remotely) on the computer system 100 to compare a nucleotide sequence with other nucleotide sequences and/or compounds stored within a data storage means. For example, the sequence comparison algorithm may compare the nucleotide sequences of the invention stored on a computer readable medium to reference sequences stored on a computer readable medium to identify homologies or structural motifs.

The parameters used with the above algorithms may be adapted depending on the sequence length and degree of homology studied. In some aspects, the parameters may be the default parameters used by the algorithms in the absence of instructions from the user. Figure 2 is a flow diagram illustrating one aspect of a process 200 for comparing a new nucleotide or protein sequence with a database of sequences in order to determine the homology levels between the new sequence and the sequences in the database. The database of sequences can be a private database stored within the computer system 100, or a public database such as GENBANK that is available through the Internet. The process 200 begins at a start state 201 and then moves to a state 202 wherein the new sequence to be compared is stored to a memory in a computer system 100. As discussed above, the memory could be any type of memory, including RAM or an internal storage device. The process 200 then moves to a state 204 wherein a database of sequences is opened for analysis and comparison. The process 200 then moves to a state 206 wherein the first sequence stored in the database is read into a memory on the computer. A comparison is then performed at a state 210 to determine if the first sequence is the same as the second sequence. It is important to note that this step is not limited to performing an exact comparison between the new sequence and the first sequence in the database. Well-known methods are known to those of skill in the art for comparing two nucleotide or protein sequences, even if they are not identical. For example, gaps can be introduced into one sequence in order to raise the homology level between the two tested sequences. The parameters that control whether gaps or other features are introduced into a sequence during comparison are normally entered by the user of the computer system. Once a comparison of the two sequences has been performed at the state 210, a determination is made at a decision state 210 whether the two sequences are the same. Of course, the term "same" is not limited to sequences that are absolutely identical. Sequences that are within the homology parameters entered by the user will be marked as "same" in the process 200. If a determination is made that the

two sequences are the same, the process 200 moves to a state 214 wherein the name of the sequence from the database is displayed to the user. This state notifies the user that the sequence with the displayed name fulfills the homology constraints that were entered. Once the name of the stored sequence is displayed to the user, the process 200 moves to a decision state 218 wherein a determination is made whether more sequences exist in the database. If no more sequences exist in the database, then the process 200 terminates at an end state 220. However, if more sequences do exist in the database, then the process 200 moves to a state 224 wherein a pointer is moved to the next sequence in the database so that it can be compared to the new sequence. In this manner, the new sequence is aligned and compared with every sequence in the database. It should be noted that if a determination had been made at the decision state 212 that the sequences were not homologous, then the process 200 would move immediately to the decision state 218 in order to determine if any other sequences were available in the database for comparison.

Accordingly, one aspect of the invention is a computer system comprising a processor, a data storage device having stored thereon a nucleic acid sequence of the invention and a sequence comparer for conducting the comparison. The sequence comparer may indicate a homology level between the sequences compared or identify structural motifs, or it may identify structural motifs in sequences which are compared to these nucleic acid codes and polypeptide codes. Figure 3 is a flow diagram illustrating one embodiment of a process 250 in a computer for determining whether two sequences are homologous. The process 250 begins at a start state 252 and then moves to a state 254 wherein a first sequence to be compared is stored to a memory. The second sequence to be compared is then stored to a memory at a state 256. The process 250 then moves to a state 260 wherein the first character in the first sequence is read and then to a state 262 wherein the first character of the second sequence is read. It should be understood that if the sequence is a nucleotide sequence, then the character would normally be either A, T, C, G or U. If the sequence is a protein sequence, then it can be a single letter amino acid code so that the first and sequence sequences can be easily compared. A determination is then made at a decision state 264 whether the two characters are the same. If they are the same, then the process 250 moves to a state 268 wherein the next characters in the first and second sequences are read. A determination is then made whether the next characters are the same. If they are, then the process 250 continues this loop until two characters are not the same. If a determination is made that the next two characters are not the same, the

process 250 moves to a decision state 274 to determine whether there are any more characters either sequence to read. If there are not any more characters to read, then the process 250 moves to a state 276 wherein the level of homology between the first and second sequences is displayed to the user. The level of homology is determined by calculating the proportion of characters between the sequences that were the same out of the total number of sequences in the first sequence. Thus, if every character in a first 100 nucleotide sequence aligned with an every character in a second sequence, the homology level would be 100%.

Alternatively, the computer program can compare a reference sequence to a sequence of the invention to determine whether the sequences differ at one or more positions. The program can record the length and identity of inserted, deleted or substituted nucleotides or amino acid residues with respect to the sequence of either the reference or the invention. The computer program may be a program which determines whether a reference sequence contains a single nucleotide polymorphism (SNP) with respect to a sequence of the invention, or, whether a sequence of the invention comprises a SNP of a known sequence. Thus, in some aspects, the computer program is a program which identifies SNPs. The method may be implemented by the computer systems described above and the method illustrated in Figure 3. The method can be performed by reading a sequence of the invention and the reference sequences through the use of the computer program and identifying differences with the computer program.

In other aspects the computer based system comprises an identifier for identifying features within a nucleic acid or polypeptide of the invention. An "identifier" refers to one or more programs which identifies certain features within a nucleic acid sequence. For example, an identifier may comprise a program which identifies an open reading frame (ORF) in a nucleic acid sequence. Figure 4 is a flow diagram illustrating one aspect of an identifier process 300 for detecting the presence of a feature in a sequence. The process 300 begins at a start state 302 and then moves to a state 304 wherein a first sequence that is to be checked for features is stored to a memory 115 in the computer system 100. The process 300 then moves to a state 306 wherein a database of sequence features is opened. Such a database would include a list of each feature's attributes along with the name of the feature. For example, a feature name could be "Initiation Codon" and the attribute would be "ATG". Another example would be the feature name "TAATAA Box" and the feature attribute would be "TAATAA". An

example of such a database is produced by the University of Wisconsin Genetics Computer Group. Alternatively, the features may be structural polypeptide motifs such as alpha helices, beta sheets, or functional polypeptide motifs such as enzymatic active sites, helix-turn-helix motifs or other motifs known to those skilled in the art. Once the database of features is opened at the state 306, the process 300 moves to a state 308 wherein the first feature is read from the database. A comparison of the attribute of the first feature with the first sequence is then made at a state 310. A determination is then made at a decision state 316 whether the attribute of the feature was found in the first sequence. If the attribute was found, then the process 300 moves to a state 318 wherein the name of the found feature is displayed to the user. The process 300 then moves to a decision state 320 wherein a determination is made whether move features exist in the database. If no more features do exist, then the process 300 terminates at an end state 324. However, if more features do exist in the database, then the process 300 reads the next sequence feature at a state 326 and loops back to the state 310 wherein the attribute of the next feature is compared against the first sequence. If the feature attribute is not found in the first sequence at the decision state 316, the process 300 moves directly to the decision state 320 in order to determine if any more features exist in the database. Thus, in one aspect, the invention provides a computer program that identifies open reading frames (ORFs).

A polypeptide or nucleic acid sequence of the invention can be stored and manipulated in a variety of data processor programs in a variety of formats. For example, a sequence can be stored as text in a word processing file, such as MicrosoftWORD or WORDPERFECT or as an ASCII file in a variety of database programs familiar to those of skill in the art, such as DB2, SYBASE, or ORACLE. In addition, many computer programs and databases may be used as sequence comparison algorithms, identifiers, or sources of reference nucleotide sequences or polypeptide sequences to be compared to a nucleic acid sequence of the invention. The programs and databases used to practice the invention include, but are not limited to: MacPattern (EMBL), DiscoveryBase (Molecular Applications Group), GeneMine (Molecular Applications Group), Look (Molecular Applications Group), MacLook (Molecular Applications Group), BLAST and BLAST2 (NCBI), BLASTN and BLASTX (Altschul et al, J. Mol. Biol. 215: 403, 1990), FASTA (Pearson and Lipman, Proc. Natl. Acad. Sci. USA, 85: 2444, 1988), FASTDB (Brutlag et al. Comp. App. Biosci. 6:237-245, 1990), Catalyst (Molecular Simulations Inc.),

Catalyst/SHAPE (Molecular Simulations Inc.), Cerius2.DBAccess (Molecular Simulations Inc.), HypoGen (Molecular Simulations Inc.), Insight II, (Molecular Simulations Inc.), Discover (Molecular Simulations Inc.), CHARMm (Molecular Simulations Inc.), Felix (Molecular Simulations Inc.), DelPhi, (Molecular Simulations Inc.), QuanteMM, (Molecular Simulations Inc.), Homology (Molecular Simulations Inc.), Modeler (Molecular Simulations Inc.), ISIS (Molecular Simulations Inc.), Quanta/Protein Design (Molecular Simulations Inc.), WebLab (Molecular Simulations Inc.), WebLab Diversity Explorer (Molecular Simulations Inc.), Gene Explorer (Molecular Simulations Inc.), SeqFold (Molecular Simulations Inc.), the MDL Available Chemicals Directory database, the MDL Drug Data Report data base, the Comprehensive Medicinal Chemistry database, Derwent's World Drug Index database, the BioByteMasterFile database, the Genbank database, and the Genseqn database. Many other programs and data bases would be apparent to one of skill in the art given the present disclosure.

Motifs which may be detected using the above programs include sequences encoding leucine zippers, helix-turn-helix motifs, glycosylation sites, ubiquitination sites, alpha helices, and beta sheets, signal sequences encoding signal peptides which direct the secretion of the encoded proteins, sequences implicated in transcription regulation such as homeoboxes, acidic stretches, enzymatic active sites, substrate binding sites, and enzymatic cleavage sites.

Hybridization of nucleic acids

The invention provides isolated or recombinant nucleic acids that hybridize under stringent conditions to an exemplary sequence of the invention, or a nucleic acid that encodes a polypeptide of the invention. The stringent conditions can be highly stringent conditions, medium stringent conditions, low stringent conditions, including the high and reduced stringency conditions described herein. In one aspect, it is the stringency of the wash conditions that set forth the conditions which determine whether a nucleic acid is within the scope of the invention, as discussed below.

In alternative embodiments, nucleic acids of the invention as defined by their ability to hybridize under stringent conditions can be between about five residues and the full length of nucleic acid of the invention; e.g., they can be at least 5, 10, 15, 20, 25, 30, 35, 40, 50, 55, 60, 65, 70, 75, 80, 90, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, or more, residues in length. Nucleic

acids shorter than full length are also included. These nucleic acids can be useful as, e.g., hybridization probes, labeling probes, PCR oligonucleotide probes, iRNA, antisense or sequences encoding antibody binding peptides (epitopes), motifs, active sites and the like.

In one aspect, nucleic acids of the invention are defined by their ability to hybridize under high stringency comprises conditions of about 50% formamide at about 37°C to 42°C. In one aspect, nucleic acids of the invention are defined by their ability to hybridize under reduced stringency comprising conditions in about 35% to 25% formamide at about 30°C to 35°C.

Alternatively, nucleic acids of the invention are defined by their ability to hybridize under high stringency comprising conditions at 42°C in 50% formamide, 5X SSPE, 0.3% SDS, and a repetitive sequence blocking nucleic acid, such as cot-1 or salmon sperm DNA (e.g., 200 n/ml sheared and denatured salmon sperm DNA). In one aspect, nucleic acids of the invention are defined by their ability to hybridize under reduced stringency conditions comprising 35% formamide at a reduced temperature of 35°C.

Following hybridization, the filter may be washed with 6X SSC, 0.5% SDS at 50°C. These conditions are considered to be "moderate" conditions above 25% formamide and "low" conditions below 25% formamide. A specific example of "moderate" hybridization conditions is when the above hybridization is conducted at 30% formamide. A specific example of "low stringency" hybridization conditions is when the above hybridization is conducted at 10% formamide.

The temperature range corresponding to a particular level of stringency can be further narrowed by calculating the purine to pyrimidine ratio of the nucleic acid of interest and adjusting the temperature accordingly. Nucleic acids of the invention are also defined by their ability to hybridize under high, medium, and low stringency conditions as set forth in Ausubel and Sambrook. Variations on the above ranges and conditions are well known in the art. Hybridization conditions are discussed further, below.

The above procedure may be modified to identify nucleic acids having decreasing levels of homology to the probe sequence. For example, to obtain nucleic acids of decreasing homology to the detectable probe, less stringent conditions may be used. For example, the hybridization temperature may be decreased in increments of 5°C from 68°C to 42°C in a hybridization buffer having a Na^+ concentration of approximately

1M. Following hybridization, the filter may be washed with 2X SSC, 0.5% SDS at the temperature of hybridization. These conditions are considered to be "moderate" conditions above 50°C and "low" conditions below 50°C. A specific example of "moderate" hybridization conditions is when the above hybridization is conducted at 55°C. A specific example of "low stringency" hybridization conditions is when the above hybridization is conducted at 45°C.

Alternatively, the hybridization may be carried out in buffers, such as 6X SSC, containing formamide at a temperature of 42°C. In this case, the concentration of formamide in the hybridization buffer may be reduced in 5% increments from 50% to 0% to identify clones having decreasing levels of homology to the probe. Following hybridization, the filter may be washed with 6X SSC, 0.5% SDS at 50°C. These conditions are considered to be "moderate" conditions above 25% formamide and "low" conditions below 25% formamide. A specific example of "moderate" hybridization conditions is when the above hybridization is conducted at 30% formamide. A specific example of "low stringency" hybridization conditions is when the above hybridization is conducted at 10% formamide.

However, the selection of a hybridization format is not critical - it is the stringency of the wash conditions that set forth the conditions which determine whether a nucleic acid is within the scope of the invention. Wash conditions used to identify nucleic acids within the scope of the invention include, e.g.: a salt concentration of about 0.02 molar at pH 7 and a temperature of at least about 50°C or about 55°C to about 60°C; or, a salt concentration of about 0.15 M NaCl at 72°C for about 15 minutes; or, a salt concentration of about 0.2X SSC at a temperature of at least about 50°C or about 55°C to about 60°C for about 15 to about 20 minutes; or, the hybridization complex is washed twice with a solution with a salt concentration of about 2X SSC containing 0.1% SDS at room temperature for 15 minutes and then washed twice by 0.1X SSC containing 0.1% SDS at 68°C for 15 minutes; or, equivalent conditions. See Sambrook, Tijssen and Ausubel for a description of SSC buffer and equivalent conditions.

These methods may be used to isolate nucleic acids of the invention.

Oligonucleotides probes and methods for using them

The invention also provides nucleic acid probes that can be used, e.g., for identifying nucleic acids encoding a polypeptide with an amylase activity or fragments thereof or for identifying amylase genes. In one aspect, the probe comprises at least 10 consecutive bases of a nucleic acid of the invention. Alternatively, a probe of the invention can be at least about 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 110, 120, 130, 150 or about 10 to 50, about 20 to 60 about 30 to 70, consecutive bases of a sequence as set forth in a nucleic acid of the invention. The probes identify a nucleic acid by binding and/or hybridization. The probes can be used in arrays of the invention, see discussion below, including, e.g., capillary arrays. The probes of the invention can also be used to isolate other nucleic acids or polypeptides.

The probes of the invention can be used to determine whether a biological sample, such as a soil sample, contains an organism having a nucleic acid sequence of the invention or an organism from which the nucleic acid was obtained. In such procedures, a biological sample potentially harboring the organism from which the nucleic acid was isolated is obtained and nucleic acids are obtained from the sample. The nucleic acids are contacted with the probe under conditions which permit the probe to specifically hybridize to any complementary sequences present in the sample. Where necessary, conditions which permit the probe to specifically hybridize to complementary sequences may be determined by placing the probe in contact with complementary sequences from samples known to contain the complementary sequence, as well as control sequences which do not contain the complementary sequence. Hybridization conditions, such as the salt concentration of the hybridization buffer, the formamide concentration of the hybridization buffer, or the hybridization temperature, may be varied to identify conditions which allow the probe to hybridize specifically to complementary nucleic acids (see discussion on specific hybridization conditions).

If the sample contains the organism from which the nucleic acid was isolated, specific hybridization of the probe is then detected. Hybridization may be detected by labeling the probe with a detectable agent such as a radioactive isotope, a fluorescent dye or an enzyme capable of catalyzing the formation of a detectable product. Many methods for using the labeled probes to detect the presence of complementary nucleic acids in a sample are familiar to those skilled in the art. These include Southern

Blots, Northern Blots, colony hybridization procedures, and dot blots. Protocols for each of these procedures are provided in Ausubel and Sambrook.

Alternatively, more than one probe (at least one of which is capable of specifically hybridizing to any complementary sequences which are present in the nucleic acid sample), may be used in an amplification reaction to determine whether the sample contains an organism containing a nucleic acid sequence of the invention (e.g., an organism from which the nucleic acid was isolated). In one aspect, the probes comprise oligonucleotides. In one aspect, the amplification reaction may comprise a PCR reaction. PCR protocols are described in Ausubel and Sambrook (see discussion on amplification reactions). In such procedures, the nucleic acids in the sample are contacted with the probes, the amplification reaction is performed, and any resulting amplification product is detected. The amplification product may be detected by performing gel electrophoresis on the reaction products and staining the gel with an intercalator such as ethidium bromide. Alternatively, one or more of the probes may be labeled with a radioactive isotope and the presence of a radioactive amplification product may be detected by autoradiography after gel electrophoresis.

Probes derived from sequences near the 3' or 5' ends of a nucleic acid sequence of the invention can also be used in chromosome walking procedures to identify clones containing additional, e.g., genomic sequences. Such methods allow the isolation of genes which encode additional proteins of interest from the host organism.

In one aspect, nucleic acid sequences of the invention are used as probes to identify and isolate related nucleic acids. In some aspects, the so-identified related nucleic acids may be cDNAs or genomic DNAs from organisms other than the one from which the nucleic acid of the invention was first isolated. In such procedures, a nucleic acid sample is contacted with the probe under conditions which permit the probe to specifically hybridize to related sequences. Hybridization of the probe to nucleic acids from the related organism is then detected using any of the methods described above.

In nucleic acid hybridization reactions, the conditions used to achieve a particular level of stringency can vary, depending on the nature of the nucleic acids being hybridized. For example, the length, degree of complementarity, nucleotide sequence composition (e.g., GC v. AT content), and nucleic acid type (e.g., RNA v. DNA) of the hybridizing regions of the nucleic acids can be considered in selecting hybridization conditions. An additional consideration is whether one of the nucleic acids is

immobilized, for example, on a filter. Hybridization can be carried out under conditions of low stringency, moderate stringency or high stringency. As an example of nucleic acid hybridization, a polymer membrane containing immobilized denatured nucleic acids is first prehybridized for 30 minutes at 45°C in a solution consisting of 0.9 M NaCl, 50 mM NaH_2PO4, pH 7.0, 5.0 mM Na_2EDTA, 0.5% SDS, 10X Denhardt's, and 0.5 mg/ml polyriboadenylic acid. Approximately 2 X 10^7 cpm (specific activity 4-9 X 10^8 cpm/ug) of ^{32}P end-labeled oligonucleotide probe can then added to the solution. After 12-16 hours of incubation, the membrane is washed for 30 minutes at room temperature (RT) in 1X SET (150 mM NaCl, 20 mM Tris hydrochloride, pH 7.8, 1 mM Na_2EDTA) containing 0.5% SDS, followed by a 30 minute wash in fresh 1X SET at Tm-10°C for the oligonucleotide probe. The membrane is then exposed to auto-radiographic film for detection of hybridization signals.

By varying the stringency of the hybridization conditions used to identify nucleic acids, such as cDNAs or genomic DNAs, which hybridize to the detectable probe, nucleic acids having different levels of homology to the probe can be identified and isolated. Stringency may be varied by conducting the hybridization at varying temperatures below the melting temperatures of the probes. The melting temperature, Tm, is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly complementary probe. Very stringent conditions are selected to be equal to or about 5°C lower than the Tm for a particular probe. The melting temperature of the probe may be calculated using the following exemplary formulas. For probes between 14 and 70 nucleotides in length the melting temperature (Tm) is calculated using the formula: Tm=81.5+16.6(log [Na+])+0.41(fraction G+C)-(600/N) where N is the length of the probe. If the hybridization is carried out in a solution containing formamide, the melting temperature may be calculated using the equation: Tm=81.5+16.6(log [Na+])+0.41(fraction G+C)-(0.63% formamide)-(600/N) where N is the length of the probe. Prehybridization may be carried out in 6X SSC, 5X Denhardt's reagent, 0.5% SDS, 100µg denatured fragmented salmon sperm DNA or 6X SSC, 5X Denhardt's reagent, 0.5% SDS, 100µg denatured fragmented salmon sperm DNA, 50% formamide. Formulas for SSC and Denhardt's and other solutions are listed, e.g., in Sambrook.

Hybridization is conducted by adding the detectable probe to the prehybridization solutions listed above. Where the probe comprises double stranded

DNA, it is denatured before addition to the hybridization solution. The filter is contacted with the hybridization solution for a sufficient period of time to allow the probe to hybridize to cDNAs or genomic DNAs containing sequences complementary thereto or homologous thereto. For probes over 200 nucleotides in length, the hybridization may be carried out at 15-25°C below the Tm. For shorter probes, such as oligonucleotide probes, the hybridization may be conducted at 5-10°C below the Tm. In one aspect, hybridizations in 6X SSC are conducted at approximately 68°C. In one aspect, hybridizations in 50% formamide containing solutions are conducted at approximately 42°C. All of the foregoing hybridizations would be considered to be under conditions of high stringency.

Following hybridization, the filter is washed to remove any non-specifically bound detectable probe. The stringency used to wash the filters can also be varied depending on the nature of the nucleic acids being hybridized, the length of the nucleic acids being hybridized, the degree of complementarity, the nucleotide sequence composition (e.g., GC v. AT content), and the nucleic acid type (e.g., RNA v. DNA). Examples of progressively higher stringency condition washes are as follows: 2X SSC, 0.1% SDS at room temperature for 15 minutes (low stringency); 0.1X SSC, 0.5% SDS at room temperature for 30 minutes to 1 hour (moderate stringency); 0.1X SSC, 0.5% SDS for 15 to 30 minutes at between the hybridization temperature and 68°C (high stringency); and 0.15M NaCl for 15 minutes at 72°C (very high stringency). A final low stringency wash can be conducted in 0.1X SSC at room temperature. The examples above are merely illustrative of one set of conditions that can be used to wash filters. One of skill in the art would know that there are numerous recipes for different stringency washes.

Nucleic acids which have hybridized to the probe can be identified by autoradiography or other conventional techniques. The above procedure may be modified to identify nucleic acids having decreasing levels of homology to the probe sequence. For example, to obtain nucleic acids of decreasing homology to the detectable probe, less stringent conditions may be used. For example, the hybridization temperature may be decreased in increments of 5°C from 68°C to 42°C in a hybridization buffer having a Na+ concentration of approximately 1M. Following hybridization, the filter may be washed with 2X SSC, 0.5% SDS at the temperature of hybridization. These conditions are considered to be "moderate" conditions above 50°C and "low" conditions below 50°C.

An example of "moderate" hybridization conditions is when the above hybridization is conducted at 55°C. An example of "low stringency" hybridization conditions is when the above hybridization is conducted at 45°C.

Alternatively, the hybridization may be carried out in buffers, such as 6X SSC, containing formamide at a temperature of 42°C. In this case, the concentration of formamide in the hybridization buffer may be reduced in 5% increments from 50% to 0% to identify clones having decreasing levels of homology to the probe. Following hybridization, the filter may be washed with 6X SSC, 0.5% SDS at 50°C. These conditions are considered to be "moderate" conditions above 25% formamide and "low" conditions below 25% formamide. A specific example of "moderate" hybridization conditions is when the above hybridization is conducted at 30% formamide. A specific example of "low stringency" hybridization conditions is when the above hybridization is conducted at 10% formamide.

These probes and methods of the invention can be used to isolate nucleic acids having a sequence with at least about 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity ("homology") to a nucleic acid sequence of the invention comprising at least about 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, 150, 200, 250, 300, 350, 400, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, or more consecutive bases thereof, and the sequences complementary thereto. Homology may be measured using an alignment algorithm, as discussed herein. For example, the homologous polynucleotides may have a coding sequence which is a naturally occurring allelic variant of one of the coding sequences described herein. Such allelic variants may have a substitution, deletion or addition of one or more nucleotides when compared to a nucleic acid of the invention.

Additionally, the probes and methods of the invention can be used to isolate nucleic acids which encode polypeptides having at least about 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99%, sequence identity (homology) to a polypeptide of the invention comprising at least 5, 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, or 150 consecutive amino acids, as

determined using a sequence alignment algorithm (e.g., such as the FASTA version 3.0t78 algorithm with the default parameters, or a BLAST 2.2.2 program with exemplary settings as set forth herein).

Inhibiting Expression of Amylase

The invention provides nucleic acids complementary to (e.g., antisense sequences to) the nucleic acid sequences of the invention. Antisense sequences are capable of inhibiting the transport, splicing or transcription of amylase-encoding genes. The inhibition can be effected through the targeting of genomic DNA or messenger RNA. The transcription or function of targeted nucleic acid can be inhibited, for example, by hybridization and/or cleavage. One particularly useful set of inhibitors provided by the present invention includes oligonucleotides which are able to either bind amylase gene or message, in either case preventing or inhibiting the production or function of amylase. The association can be through sequence specific hybridization. Another useful class of inhibitors includes oligonucleotides which cause inactivation or cleavage of amylase message. The oligonucleotide can have enzyme activity which causes such cleavage, such as ribozymes. The oligonucleotide can be chemically modified or conjugated to an enzyme or composition capable of cleaving the complementary nucleic acid. A pool of many different such oligonucleotides can be screened for those with the desired activity.

Antisense Oligonucleotides

The invention provides antisense oligonucleotides capable of binding amylase message which can inhibit proteolytic activity by targeting mRNA. Strategies for designing antisense oligonucleotides are well described in the scientific and patent literature, and the skilled artisan can design such amylase oligonucleotides using the novel reagents of the invention. For example, gene walking/ RNA mapping protocols to screen for effective antisense oligonucleotides are well known in the art, see, e.g., Ho (2000) Methods Enzymol. 314:168-183, describing an RNA mapping assay, which is based on standard molecular techniques to provide an easy and reliable method for potent antisense sequence selection. See also Smith (2000) Eur. J. Pharm. Sci. 11:191-198.

Naturally occurring nucleic acids are used as antisense oligonucleotides. The antisense oligonucleotides can be of any length; for example, in alternative aspects, the antisense oligonucleotides are between about 5 to 100, about 10 to 80, about 15 to 60, about 18 to 40. The optimal length can be determined by routine screening. The

antisense oligonucleotides can be present at any concentration. The optimal concentration can be determined by routine screening. A wide variety of synthetic, non-naturally occurring nucleotide and nucleic acid analogues are known which can address this potential problem. For example, peptide nucleic acids (PNAs) containing non-ionic backbones, such as N-(2-aminoethyl) glycine units can be used. Antisense oligonucleotides having phosphorothioate linkages can also be used, as described in WO 97/03211; WO 96/39154; Mata (1997) Toxicol Appl Pharmacol 144:189-197; Antisense Therapeutics, ed. Agrawal (Humana Press, Totowa, N.J., 1996). Antisense oligonucleotides having synthetic DNA backbone analogues provided by the invention can also include phosphoro-dithioate, methylphosphonate, phosphoramidate, alkyl phosphotriester, sulfamate, 3'-thioacetal, methylene(methylimino), 3'-N-carbamate, and morpholino carbamate nucleic acids, as described above.

Combinatorial chemistry methodology can be used to create vast numbers of oligonucleotides that can be rapidly screened for specific oligonucleotides that have appropriate binding affinities and specificities toward any target, such as the sense and antisense amylase sequences of the invention (see, e.g., Gold (1995) J. of Biol. Chem. 270:13581-13584).

Inhibitory Ribozymes

The invention provides ribozymes capable of binding amylase message. These ribozymes can inhibit amylase activity by, e.g., targeting mRNA. Strategies for designing ribozymes and selecting the amylase-specific antisense sequence for targeting are well described in the scientific and patent literature, and the skilled artisan can design such ribozymes using the novel reagents of the invention. Ribozymes act by binding to a target RNA through the target RNA binding portion of a ribozyme which is held in close proximity to an enzymatic portion of the RNA that cleaves the target RNA. Thus, the ribozyme recognizes and binds a target RNA through complementary base-pairing, and once bound to the correct site, acts enzymatically to cleave and inactivate the target RNA. Cleavage of a target RNA in such a manner will destroy its ability to direct synthesis of an encoded protein if the cleavage occurs in the coding sequence. After a ribozyme has bound and cleaved its RNA target, it can be released from that RNA to bind and cleave new targets repeatedly.

In some circumstances, the enzymatic nature of a ribozyme can be advantageous over other technologies, such as antisense technology (where a nucleic acid molecule simply binds to a nucleic acid target to block its transcription, translation or association with another molecule) as the effective concentration of ribozyme necessary to effect a therapeutic treatment can be lower than that of an antisense oligonucleotide. This potential advantage reflects the ability of the ribozyme to act enzymatically. Thus, a single ribozyme molecule is able to cleave many molecules of target RNA. In addition, a ribozyme is typically a highly specific inhibitor, with the specificity of inhibition depending not only on the base pairing mechanism of binding, but also on the mechanism by which the molecule inhibits the expression of the RNA to which it binds. That is, the inhibition is caused by cleavage of the RNA target and so specificity is defined as the ratio of the rate of cleavage of the targeted RNA over the rate of cleavage of non-targeted RNA. This cleavage mechanism is dependent upon factors additional to those involved in base pairing. Thus, the specificity of action of a ribozyme can be greater than that of antisense oligonucleotide binding the same RNA site.

The ribozyme of the invention, e.g., an enzymatic ribozyme RNA molecule, can be formed in a hammerhead motif, a hairpin motif, as a hepatitis delta virus motif, a group I intron motif and/or an RNaseP-like RNA in association with an RNA guide sequence. Examples of hammerhead motifs are described by, e.g., Rossi (1992) Aids Research and Human Retroviruses 8:183; hairpin motifs by Hampel (1989) Biochemistry 28:4929, and Hampel (1990) Nuc. Acids Res. 18:299; the hepatitis delta virus motif by Perrotta (1992) Biochemistry 31:16; the RNaseP motif by Guerrier-Takada (1983) Cell 35:849; and the group I intron by Cech U.S. Pat. No. 4,987,071. The recitation of these specific motifs is not intended to be limiting. Those skilled in the art will recognize that a ribozyme of the invention, e.g., an enzymatic RNA molecule of this invention, can have a specific substrate binding site complementary to one or more of the target gene RNA regions. A ribozyme of the invention can have a nucleotide sequence within or surrounding that substrate binding site which imparts an RNA cleaving activity to the molecule.

RNA interference (RNAi)

In one aspect, the invention provides an RNA inhibitory molecule, a so-called "RNAi" molecule, comprising an amylase sequence of the invention. The RNAi

molecule comprises a double-stranded RNA (dsRNA) molecule. The RNAi can inhibit expression of an amylase gene. In one aspect, the RNAi is about 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 or more duplex nucleotides in length. While the invention is not limited by any particular mechanism of action, the RNAi can enter a cell and cause the degradation of a single-stranded RNA (ssRNA) of similar or identical sequences, including endogenous mRNAs. When a cell is exposed to double-stranded RNA (dsRNA), mRNA from the homologous gene is selectively degraded by a process called RNA interference (RNAi). A possible basic mechanism behind RNAi is the breaking of a double-stranded RNA (dsRNA) matching a specific gene sequence into short pieces called short interfering RNA, which trigger the degradation of mRNA that matches its sequence. In one aspect, the RNAi's of the invention are used in gene-silencing therapeutics, see, e.g., Shuey (2002) Drug Discov. Today 7:1040-1046. In one aspect, the invention provides methods to selectively degrade RNA using the RNAi's of the invention. The process may be practiced *in vitro*, *ex vivo* or *in vivo*. In one aspect, the RNAi molecules of the invention can be used to generate a loss-of-function mutation in a cell, an organ or an animal. Methods for making and using RNAi molecules for selectively degrade RNA are well known in the art, see, e.g., U.S. Patent No. 6,506,559; 6,511,824; 6,515,109; 6,489,127.

Modification of Nucleic Acids

The invention provides methods of generating variants of the nucleic acids of the invention, e.g., those encoding an amylase. These methods can be repeated or used in various combinations to generate amylases having an altered or different activity or an altered or different stability from that of an amylase encoded by the template nucleic acid. These methods also can be repeated or used in various combinations, e.g., to generate variations in gene/ message expression, message translation or message stability. In another aspect, the genetic composition of a cell is altered by, e.g., modification of a homologous gene ex vivo, followed by its reinsertion into the cell.

A nucleic acid of the invention can be altered by any means. For example, random or stochastic methods, or, non-stochastic, or "directed evolution," methods, see, e.g., U.S. Patent No. 6,361,974. Methods for random mutation of genes are well known in the art, see, e.g., U.S. Patent No. 5,830,696. For example, mutagens can be used to randomly mutate a gene. Mutagens include, e.g., ultraviolet light or gamma irradiation,

or a chemical mutagen, e.g., mitomycin, nitrous acid, photoactivated psoralens, alone or in combination, to induce DNA breaks amenable to repair by recombination. Other chemical mutagens include, for example, sodium bisulfite, nitrous acid, hydroxylamine, hydrazine or formic acid. Other mutagens are analogues of nucleotide precursors, e.g., nitrosoguanidine, 5-bromouracil, 2-aminopurine, or acridine. These agents can be added to a PCR reaction in place of the nucleotide precursor thereby mutating the sequence. Intercalating agents such as proflavine, acriflavine, quinacrine and the like can also be used.

Any technique in molecular biology can be used, e.g., random PCR mutagenesis, see, e.g., Rice (1992) Proc. Natl. Acad. Sci. USA 89:5467-5471; or, combinatorial multiple cassette mutagenesis, see, e.g., Crameri (1995) Biotechniques 18:194-196. Alternatively, nucleic acids, e.g., genes, can be reassembled after random, or "stochastic," fragmentation, see, e.g., U.S. Patent Nos. 6,291,242; 6,287,862; 6,287,861; 5,955,358; 5,830,721; 5,824,514; 5,811,238; 5,605,793. In alternative aspects, modifications, additions or deletions are introduced by error-prone PCR, shuffling, oligonucleotide-directed mutagenesis, assembly PCR, sexual PCR mutagenesis, in vivo mutagenesis, cassette mutagenesis, recursive ensemble mutagenesis, exponential ensemble mutagenesis, site-specific mutagenesis, gene reassembly, gene site saturated mutagenesis (GSSM), synthetic ligation reassembly (SLR), recombination, recursive sequence recombination, phosphothioate-modified DNA mutagenesis, uracil-containing template mutagenesis, gapped duplex mutagenesis, point mismatch repair mutagenesis, repair-deficient host strain mutagenesis, chemical mutagenesis, radiogenic mutagenesis, deletion mutagenesis, restriction-selection mutagenesis, restriction-purification mutagenesis, artificial gene synthesis, ensemble mutagenesis, chimeric nucleic acid multimer creation, and/or a combination of these and other methods.

The following publications describe a variety of recursive recombination procedures and/or methods which can be incorporated into the methods of the invention: Stemmer (1999) "Molecular breeding of viruses for targeting and other clinical properties" Tumor Targeting 4:1-4; Ness (1999) Nature Biotechnology 17:893-896; Chang (1999) "Evolution of a cytokine using DNA family shuffling" Nature Biotechnology 17:793-797; Minshull (1999) "Protein evolution by molecular breeding" Current Opinion in Chemical Biology 3:284-290; Christians (1999) "Directed evolution of thymidine kinase for AZT phosphorylation using DNA family shuffling" Nature

Biotechnology 17:259-264; Crameri (1998) "DNA shuffling of a family of genes from diverse species accelerates directed evolution" Nature 391:288-291; Crameri (1997) "Molecular evolution of an arsenate detoxification pathway by DNA shuffling," Nature Biotechnology 15:436-438; Zhang (1997) "Directed evolution of an effective fucosidase from a galactosidase by DNA shuffling and screening" Proc. Natl. Acad. Sci. USA 94:4504-4509; Patten et al. (1997) "Applications of DNA Shuffling to Pharmaceuticals and Vaccines" Current Opinion in Biotechnology 8:724-733; Crameri et al. (1996) "Construction and evolution of antibody-phage libraries by DNA shuffling" Nature Medicine 2:100-103; Gates et al. (1996) "Affinity selective isolation of ligands from peptide libraries through display on a lac repressor `headpiece dimer`" Journal of Molecular Biology 255:373-386; Stemmer (1996) "Sexual PCR and Assembly PCR" In: The Encyclopedia of Molecular Biology. VÇH Publishers, New York. pp.447-457; Crameri and Stemmer (1995) "Combinatorial multiple cassette mutagenesis creates all the permutations of mutant and wildtype cassettes" BioTechniques 18:194-195; Stemmer et al. (1995) "Single-step assembly of a gene and entire plasmid form large numbers of oligodeoxyribonucleotides" Gene, 164:49-53; Stemmer (1995) "The Evolution of Molecular Computation" Science 270: 1510; Stemmer (1995) "Searching Sequence Space" Bio/Technology 13:549-553; Stemmer (1994) "Rapid evolution of a protein in vitro by DNA shuffling" Nature 370:389-391; and Stemmer (1994) "DNA shuffling by random fragmentation and reassembly: In vitro recombination for molecular evolution." Proc. Natl. Acad. Sci. USA 91:10747-10751.

Mutational methods of generating diversity include, for example, site-directed mutagenesis (Ling et al. (1997) "Approaches to DNA mutagenesis: an overview" Anal Biochem. 254(2): 157-178; Dale et al. (1996) "Oligonucleotide-directed random mutagenesis using the phosphorothioate method" Methods Mol. Biol. 57:369-374; Smith (1985) "In vitro mutagenesis" Ann. Rev. Genet. 19:423-462; Botstein & Shortle (1985) "Strategies and applications of in vitro mutagenesis" Science 229:1193-1201; Carter (1986) "Site-directed mutagenesis" Biochem. J. 237:1-7; and Kunkel (1987) "The efficiency of oligonucleotide directed mutagenesis" in Nucleic Acids & Molecular Biology (Eckstein, F. and Lilley, D. M. J. eds., Springer Verlag, Berlin)); mutagenesis using uracil containing templates (Kunkel (1985) "Rapid and efficient site-specific mutagenesis without phenotypic selection" Proc. Natl. Acad. Sci. USA 82:488-492; Kunkel et al. (1987) "Rapid and efficient site-specific mutagenesis without phenotypic

selection" Methods in Enzymol. 154, 367-382; and Bass et al. (1988) "Mutant Trp repressors with new DNA-binding specificities" Science 242:240-245); oligonucleotide-directed mutagenesis (Methods in Enzymol. 100: 468-500 (1983); Methods in Enzymol. 154: 329-350 (1987); Zoller & Smith (1982) "Oligonucleotide-directed mutagenesis using M13-derived vectors: an efficient and general procedure for the production of point mutations in any DNA fragment" Nucleic Acids Res. 10:6487-6500; Zoller & Smith (1983) "Oligonucleotide-directed mutagenesis of DNA fragments cloned into M13 vectors" Methods in Enzymol. 100:468-500; and Zoller & Smith (1987) Oligonucleotide-directed mutagenesis: a simple method using two oligonucleotide primers and a single-stranded DNA template" Methods in Enzymol. 154:329-350); phosphorothioate-modified DNA mutagenesis (Taylor et al. (1985) "The use of phosphorothioate-modified DNA in restriction enzyme reactions to prepare nicked DNA" Nucl. Acids Res. 13: 8749-8764; Taylor et al. (1985) "The rapid generation of oligonucleotide-directed mutations at high frequency using phosphorothioate-modified DNA" Nucl. Acids Res. 13: 8765-8787 (1985); Nakamaye (1986) "Inhibition of restriction endonuclease Nci I cleavage by phosphorothioate groups and its application to oligonucleotide-directed mutagenesis" Nucl. Acids Res. 14: 9679-9698; Sayers et al. (1988) "Y-T Exonucleases in phosphorothioate-based oligonucleotide-directed mutagenesis" Nucl. Acids Res. 16:791-802; and Sayers et al. (1988) "Strand specific cleavage of phosphorothioate-containing DNA by reaction with restriction endonucleases in the presence of ethidium bromide" Nucl. Acids Res. 16: 803-814); mutagenesis using gapped duplex DNA (Kramer et al. (1984) "The gapped duplex DNA approach to oligonucleotide-directed mutation construction" Nucl. Acids Res. 12: 9441-9456; Kramer & Fritz (1987) Methods in Enzymol. "Oligonucleotide-directed construction of mutations via gapped duplex DNA" 154:350-367; Kramer et al. (1988) "Improved enzymatic in vitro reactions in the gapped duplex DNA approach to oligonucleotide-directed construction of mutations" Nucl. Acids Res. 16: 7207; and Fritz et al. (1988) "Oligonucleotide-directed construction of mutations: a gapped duplex DNA procedure without enzymatic reactions *in vitro*" Nucl. Acids Res. 16: 6987-6999).

Additional protocols that can be used to practice the invention include point mismatch repair (Kramer (1984) "Point Mismatch Repair" Cell 38:879-887), mutagenesis using repair-deficient host strains (Carter et al. (1985) "Improved oligonucleotide site-directed mutagenesis using M13 vectors" Nucl. Acids Res. 13: 4431-

4443; and Carter (1987) "Improved oligonucleotide-directed mutagenesis using M13 vectors" Methods in Enzymol. 154: 382-403), deletion mutagenesis (Eghtedarzadeh (1986) "Use of oligonucleotides to generate large deletions" Nucl. Acids Res. 14: 5115), restriction-selection and restriction-selection and restriction-purification (Wells et al. (1986) "Importance of hydrogen-bond formation in stabilizing the transition state of subtilisin" Phil. Trans. R. Soc. Lond. A 317: 415-423), mutagenesis by total gene synthesis (Nambiar et al. (1984) "Total synthesis and cloning of a gene coding for the ribonuclease S protein" Science 223: 1299-1301; Sakamar and Khorana (1988) "Total synthesis and expression of a gene for the a-subunit of bovine rod outer segment guanine nucleotide-binding protein (transducin)" Nucl. Acids Res. 14: 6361-6372; Wells et al. (1985) "Cassette mutagenesis: an efficient method for generation of multiple mutations at defined sites" Gene 34:315-323; and Grundstrom et al. (1985) "Oligonucleotide-directed mutagenesis by microscale `shot-gun` gene synthesis" Nucl. Acids Res. 13: 3305-3316), double-strand break repair (Mandecki (1986); Arnold (1993) "Protein engineering for unusual environments" Current Opinion in Biotechnology 4:450-455. "Oligonucleotide-directed double-strand break repair in plasmids of Escherichia coli: a method for site-specific mutagenesis" Proc. Natl. Acad. Sci. USA, 83:7177-7181). Additional details on many of the above methods can be found in Methods in Enzymology Volume 154, which also describes useful controls for trouble-shooting problems with various mutagenesis methods.

Protocols that can be used to practice the invention are described, e.g., in U.S. Patent Nos. 5,605,793 to Stemmer (Feb. 25, 1997), "Methods for In Vitro Recombination;" U.S. Pat. No. 5,811,238 to Stemmer et al. (Sep. 22, 1998) "Methods for Generating Polynucleotides having Desired Characteristics by Iterative Selection and Recombination;" U.S. Pat. No. 5,830,721 to Stemmer et al. (Nov. 3, 1998), "DNA Mutagenesis by Random Fragmentation and Reassembly;" U.S. Pat. No. 5,834,252 to Stemmer, et al. (Nov. 10, 1998) "End-Complementary Polymerase Reaction;" U.S. Pat. No. 5,837,458 to Minshull, et al. (Nov. 17, 1998), "Methods and Compositions for Cellular and Metabolic Engineering;" WO 95/22625, Stemmer and Crameri, "Mutagenesis by Random Fragmentation and Reassembly;" WO 96/33207 by Stemmer and Lipschutz "End Complementary Polymerase Chain Reaction;" WO 97/20078 by Stemmer and Crameri "Methods for Generating Polynucleotides having Desired Characteristics by Iterative Selection and Recombination;" WO 97/35966 by Minshull

and Stemmer, "Methods and Compositions for Cellular and Metabolic Engineering;" WO 99/41402 by Punnonen et al. "Targeting of Genetic Vaccine Vectors;" WO 99/41383 by Punnonen et al. "Antigen Library Immunization;" WO 99/41369 by Punnonen et al. "Genetic Vaccine Vector Engineering;" WO 99/41368 by Punnonen et al. "Optimization of Immunomodulatory Properties of Genetic Vaccines;" EP 752008 by Stemmer and Crameri, "DNA Mutagenesis by Random Fragmentation and Reassembly;" EP 0932670 by Stemmer "Evolving Cellular DNA Uptake by Recursive Sequence Recombination;" WO 99/23107 by Stemmer et al., "Modification of Virus Tropism and Host Range by Viral Genome Shuffling;" WO 99/21979 by Apt et al., "Human Papillomavirus Vectors;" WO 98/31837 by del Cardayre et al. "Evolution of Whole Cells and Organisms by Recursive Sequence Recombination;" WO 98/27230 by Patten and Stemmer, "Methods and Compositions for Polypeptide Engineering;" WO 98/27230 by Stemmer et al., "Methods for Optimization of Gene Therapy by Recursive Sequence Shuffling and Selection," WO 00/00632, "Methods for Generating Highly Diverse Libraries," WO 00/09679, "Methods for Obtaining in Vitro Recombined Polynucleotide Sequence Banks and Resulting Sequences," WO 98/42832 by Arnold et al., "Recombination of Polynucleotide Sequences Using Random or Defined Primers," WO 99/29902 by Arnold et al., "Method for Creating Polynucleotide and Polypeptide Sequences," WO 98/41653 by Vind, "An in Vitro Method for Construction of a DNA Library," WO 98/41622 by Borchert et al., "Method for Constructing a Library Using DNA Shuffling," and WO 98/42727 by Pati and Zarling, "Sequence Alterations using Homologous Recombination."

Protocols that can be used to practice the invention (providing details regarding various diversity generating methods) are described, e.g., in U.S. Patent application serial no. (USSN) 09/407,800, "SHUFFLING OF CODON ALTERED GENES" by Patten et al. filed Sep. 28, 1999; "EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE SEQUENCE RECOMBINATION" by del Cardayre et al., United States Patent No. 6,379,964; "OLIGONUCLEOTIDE MEDIATED NUCLEIC ACID RECOMBINATION" by Crameri et al., United States Patent Nos. 6,319,714; 6,368,861; 6,376,246; 6,423,542; 6,426,224 and PCT/US00/01203; "USE OF CODON-VARIED OLIGONUCLEOTIDE SYNTHESIS FOR SYNTHETIC SHUFFLING" by Welch et al., United States Patent No. 6,436,675; "METHODS FOR MAKING CHARACTER STRINGS, POLYNUCLEOTIDES & POLYPEPTIDES HAVING DESIRED CHARACTERISTICS" by Selifonov et al., filed Jan. 18, 2000,

(PCT/US00/01202) and, e.g. "METHODS FOR MAKING CHARACTER STRINGS, POLYNUCLEOTIDES & POLYPEPTIDES HAVING DESIRED CHARACTERISTICS" by Selifonov et al., filed Jul. 18, 2000 (U.S. Ser. No. 09/618,579); "METHODS OF POPULATING DATA STRUCTURES FOR USE IN EVOLUTIONARY SIMULATIONS" by Selifonov and Stemmer, filed Jan. 18, 2000 (PCT/US00/01138); and "SINGLE-STRANDED NUCLEIC ACID TEMPLATE-MEDIATED RECOMBINATION AND NUCLEIC ACID FRAGMENT ISOLATION" by Affholter, filed Sep. 6, 2000 (U.S. Ser. No. 09/656,549); and United States Patent Nos. 6,177,263; 6,153,410.

Non-stochastic, or "directed evolution," methods include, e.g., gene site saturation mutagenesis (GSSM™), synthetic ligation reassembly (SLR), or a combination thereof are used to modify the nucleic acids of the invention to generate amylases with new or altered properties (e.g., activity under highly acidic or alkaline conditions, high temperatures, and the like). Polypeptides encoded by the modified nucleic acids can be screened for an activity before testing for proteolytic or other activity. Any testing modality or protocol can be used, e.g., using a capillary array platform. See, e.g., U.S. Patent Nos. 6,361,974; 6,280,926; 5,939,250.

Saturation mutagenesis, or, GSSM™

In one aspect, codon primers containing a degenerate N,N,G/T sequence are used to introduce point mutations into a polynucleotide, e.g., an amylase or an antibody of the invention, so as to generate a set of progeny polypeptides in which a full range of single amino acid substitutions is represented at each amino acid position, e.g., an amino acid residue in an enzyme active site or ligand binding site targeted to be modified. These oligonucleotides can comprise a contiguous first homologous sequence, a degenerate N,N,G/T sequence, and, optionally, a second homologous sequence. The downstream progeny translational products from the use of such oligonucleotides include all possible amino acid changes at each amino acid site along the polypeptide, because the degeneracy of the N,N,G/T sequence includes codons for all 20 amino acids. In one aspect, one such degenerate oligonucleotide (comprised of, e.g., one degenerate N,N,G/T cassette) is used for subjecting each original codon in a parental polynucleotide template to a full range of codon substitutions. In another aspect, at least two degenerate cassettes are used – either in the same oligonucleotide or not, for subjecting at least two original

codons in a parental polynucleotide template to a full range of codon substitutions. For example, more than one N,N,G/T sequence can be contained in one oligonucleotide to introduce amino acid mutations at more than one site. This plurality of N,N,G/T sequences can be directly contiguous, or separated by one or more additional nucleotide sequence(s). In another aspect, oligonucleotides serviceable for introducing additions and deletions can be used either alone or in combination with the codons containing an N,N,G/T sequence, to introduce any combination or permutation of amino acid additions, deletions, and/or substitutions.

In one aspect, simultaneous mutagenesis of two or more contiguous amino acid positions is done using an oligonucleotide that contains contiguous N,N,G/T triplets, i.e. a degenerate (N,N,G/T)n sequence. In another aspect, degenerate cassettes having less degeneracy than the N,N,G/T sequence are used. For example, it may be desirable in some instances to use (e.g. in an oligonucleotide) a degenerate triplet sequence comprised of only one N, where said N can be in the first second or third position of the triplet. Any other bases including any combinations and permutations thereof can be used in the remaining two positions of the triplet. Alternatively, it may be desirable in some instances to use (e.g. in an oligo) a degenerate N,N,N triplet sequence.

In one aspect, use of degenerate triplets (e.g., N,N,G/T triplets) allows for systematic and easy generation of a full range of possible natural amino acids (for a total of 20 amino acids) into each and every amino acid position in a polypeptide (in alternative aspects, the methods also include generation of less than all possible substitutions per amino acid residue, or codon, position). For example, for a 100 amino acid polypeptide, 2000 distinct species (i.e. 20 possible amino acids per position X 100 amino acid positions) can be generated. Through the use of an oligonucleotide or set of oligonucleotides containing a degenerate N,N,G/T triplet, 32 individual sequences can code for all 20 possible natural amino acids. Thus, in a reaction vessel in which a parental polynucleotide sequence is subjected to saturation mutagenesis using at least one such oligonucleotide, there are generated 32 distinct progeny polynucleotides encoding 20 distinct polypeptides. In contrast, the use of a non-degenerate oligonucleotide in site-directed mutagenesis leads to only one progeny polypeptide product per reaction vessel. Nondegenerate oligonucleotides can optionally be used in combination with degenerate primers disclosed; for example, nondegenerate oligonucleotides can be used to generate specific point mutations in a working polynucleotide. This provides one means to

generate specific silent point mutations, point mutations leading to corresponding amino acid changes, and point mutations that cause the generation of stop codons and the corresponding expression of polypeptide fragments.

In one aspect, each saturation mutagenesis reaction vessel contains polynucleotides encoding at least 20 progeny polypeptide (e.g., amylases) molecules such that all 20 natural amino acids are represented at the one specific amino acid position corresponding to the codon position mutagenized in the parental polynucleotide (other aspects use less than all 20 natural combinations). The 32-fold degenerate progeny polypeptides generated from each saturation mutagenesis reaction vessel can be subjected to clonal amplification (e.g. cloned into a suitable host, e.g., E. coli host, using, e.g., an expression vector) and subjected to expression screening. When an individual progeny polypeptide is identified by screening to display a favorable change in property (when compared to the parental polypeptide, such as increased proteolytic activity under alkaline or acidic conditions), it can be sequenced to identify the correspondingly favorable amino acid substitution contained therein.

In one aspect, upon mutagenizing each and every amino acid position in a parental polypeptide using saturation mutagenesis as disclosed herein, favorable amino acid changes may be identified at more than one amino acid position. One or more new progeny molecules can be generated that contain a combination of all or part of these favorable amino acid substitutions. For example, if 2 specific favorable amino acid changes are identified in each of 3 amino acid positions in a polypeptide, the permutations include 3 possibilities at each position (no change from the original amino acid, and each of two favorable changes) and 3 positions. Thus, there are 3 x 3 x 3 or 27 total possibilities, including 7 that were previously examined - 6 single point mutations (i.e. 2 at each of three positions) and no change at any position.

In another aspect, site-saturation mutagenesis can be used together with another stochastic or non-stochastic means to vary sequence, e.g., synthetic ligation reassembly (see below), shuffling, chimerization, recombination and other mutagenizing processes and mutagenizing agents. This invention provides for the use of any mutagenizing process(es), including saturation mutagenesis, in an iterative manner.

Synthetic Ligation Reassembly (SLR)

The invention provides a non-stochastic gene modification system termed

"synthetic ligation reassembly," or simply "SLR," a "directed evolution process," to generate polypeptides, e.g., amylases or antibodies of the invention, with new or altered properties. SLR is a method of ligating oligonucleotide fragments together non-stochastically. This method differs from stochastic oligonucleotide shuffling in that the nucleic acid building blocks are not shuffled, concatenated or chimerized randomly, but rather are assembled non-stochastically. See, e.g., U.S. Patent Application Serial No. (USSN) 09/332,835 entitled "Synthetic Ligation Reassembly in Directed Evolution" and filed on June 14, 1999 ("USSN 09/332,835"). In one aspect, SLR comprises the following steps: (a) providing a template polynucleotide, wherein the template polynucleotide comprises sequence encoding a homologous gene; (b) providing a plurality of building block polynucleotides, wherein the building block polynucleotides are designed to cross-over reassemble with the template polynucleotide at a predetermined sequence, and a building block polynucleotide comprises a sequence that is a variant of the homologous gene and a sequence homologous to the template polynucleotide flanking the variant sequence; (c) combining a building block polynucleotide with a template polynucleotide such that the building block polynucleotide cross-over reassembles with the template polynucleotide to generate polynucleotides comprising homologous gene sequence variations.

SLR does not depend on the presence of high levels of homology between polynucleotides to be rearranged. Thus, this method can be used to non-stochastically generate libraries (or sets) of progeny molecules comprised of over 10^{100} different chimeras. SLR can be used to generate libraries comprised of over 10^{1000} different progeny chimeras. Thus, aspects of the present invention include non-stochastic methods of producing a set of finalized chimeric nucleic acid molecule shaving an overall assembly order that is chosen by design. This method includes the steps of generating by design a plurality of specific nucleic acid building blocks having serviceable mutually compatible ligatable ends, and assembling these nucleic acid building blocks, such that a designed overall assembly order is achieved.

The mutually compatible ligatable ends of the nucleic acid building blocks to be assembled are considered to be "serviceable" for this type of ordered assembly if they enable the building blocks to be coupled in predetermined orders. Thus, the overall assembly order in which the nucleic acid building blocks can be coupled is specified by the design of the ligatable ends. If more than one assembly step is to be used, then the

overall assembly order in which the nucleic acid building blocks can be coupled is also specified by the sequential order of the assembly step(s). In one aspect, the annealed building pieces are treated with an enzyme, such as a ligase (e.g. T4 DNA ligase), to achieve covalent bonding of the building pieces.

In one aspect, the design of the oligonucleotide building blocks is obtained by analyzing a set of progenitor nucleic acid sequence templates that serve as a basis for producing a progeny set of finalized chimeric polynucleotides. These parental oligonucleotide templates thus serve as a source of sequence information that aids in the design of the nucleic acid building blocks that are to be mutagenized, e.g., chimerized or shuffled. In one aspect of this method, the sequences of a plurality of parental nucleic acid templates are aligned in order to select one or more demarcation points. The demarcation points can be located at an area of homology, and are comprised of one or more nucleotides. These demarcation points are preferably shared by at least two of the progenitor templates. The demarcation points can thereby be used to delineate the boundaries of oligonucleotide building blocks to be generated in order to rearrange the parental polynucleotides. The demarcation points identified and selected in the progenitor molecules serve as potential chimerization points in the assembly of the final chimeric progeny molecules. A demarcation point can be an area of homology (comprised of at least one homologous nucleotide base) shared by at least two parental polynucleotide sequences. Alternatively, a demarcation point can be an area of homology that is shared by at least half of the parental polynucleotide sequences, or, it can be an area of homology that is shared by at least two thirds of the parental polynucleotide sequences. Even more preferably a serviceable demarcation points is an area of homology that is shared by at least three fourths of the parental polynucleotide sequences, or, it can be shared by at almost all of the parental polynucleotide sequences. In one aspect, a demarcation point is an area of homology that is shared by all of the parental polynucleotide sequences.

In one aspect, a ligation reassembly process is performed exhaustively in order to generate an exhaustive library of progeny chimeric polynucleotides. In other words, all possible ordered combinations of the nucleic acid building blocks are represented in the set of finalized chimeric nucleic acid molecules. At the same time, in another aspect, the assembly order (i.e. the order of assembly of each building block in the 5' to 3 sequence of each finalized chimeric nucleic acid) in each combination is by

design (or non-stochastic) as described above. Because of the non-stochastic nature of this invention, the possibility of unwanted side products is greatly reduced.

In another aspect, the ligation reassembly method is performed systematically. For example, the method is performed in order to generate a systematically compartmentalized library of progeny molecules, with compartments that can be screened systematically, e.g. one by one. In other words this invention provides that, through the selective and judicious use of specific nucleic acid building blocks, coupled with the selective and judicious use of sequentially stepped assembly reactions, a design can be achieved where specific sets of progeny products are made in each of several reaction vessels. This allows a systematic examination and screening procedure to be performed. Thus, these methods allow a potentially very large number of progeny molecules to be examined systematically in smaller groups. Because of its ability to perform chimerizations in a manner that is highly flexible yet exhaustive and systematic as well, particularly when there is a low level of homology among the progenitor molecules, these methods provide for the generation of a library (or set) comprised of a large number of progeny molecules. Because of the non-stochastic nature of the instant ligation reassembly invention, the progeny molecules generated preferably comprise a library of finalized chimeric nucleic acid molecules having an overall assembly order that is chosen by design. The saturation mutagenesis and optimized directed evolution methods also can be used to generate different progeny molecular species. It is appreciated that the invention provides freedom of choice and control regarding the selection of demarcation points, the size and number of the nucleic acid building blocks, and the size and design of the couplings. It is appreciated, furthermore, that the requirement for intermolecular homology is highly relaxed for the operability of this invention. In fact, demarcation points can even be chosen in areas of little or no intermolecular homology. For example, because of codon wobble, i.e. the degeneracy of codons, nucleotide substitutions can be introduced into nucleic acid building blocks without altering the amino acid originally encoded in the corresponding progenitor template. Alternatively, a codon can be altered such that the coding for an originally amino acid is altered. This invention provides that such substitutions can be introduced into the nucleic acid building block in order to increase the incidence of intermolecular homologous demarcation points and thus to allow an increased number of couplings to be

achieved among the building blocks, which in turn allows a greater number of progeny chimeric molecules to be generated.

In another aspect, the synthetic nature of the step in which the building blocks are generated allows the design and introduction of nucleotides (e.g., one or more nucleotides, which may be, for example, codons or introns or regulatory sequences) that can later be optionally removed in an in vitro process (e.g. by mutagenesis) or in an in vivo process (e.g. by utilizing the gene splicing ability of a host organism). It is appreciated that in many instances the introduction of these nucleotides may also be desirable for many other reasons in addition to the potential benefit of creating a serviceable demarcation point.

In one aspect, a nucleic acid building block is used to introduce an intron. Thus, functional introns are introduced into a man-made gene manufactured according to the methods described herein. The artificially introduced intron(s) can be functional in a host cells for gene splicing much in the way that naturally-occurring introns serve functionally in gene splicing.

Optimized Directed Evolution System

The invention provides a non-stochastic gene modification system termed "optimized directed evolution system" to generate polypeptides, e.g., amylases or antibodies of the invention, with new or altered properties. Optimized directed evolution is directed to the use of repeated cycles of reductive reassortment, recombination and selection that allow for the directed molecular evolution of nucleic acids through recombination. Optimized directed evolution allows generation of a large population of evolved chimeric sequences, wherein the generated population is significantly enriched for sequences that have a predetermined number of crossover events.

A crossover event is a point in a chimeric sequence where a shift in sequence occurs from one parental variant to another parental variant. Such a point is normally at the juncture of where oligonucleotides from two parents are ligated together to form a single sequence. This method allows calculation of the correct concentrations of oligonucleotide sequences so that the final chimeric population of sequences is enriched for the chosen number of crossover events. This provides more control over choosing chimeric variants having a predetermined number of crossover events.

In addition, this method provides a convenient means for exploring a tremendous amount of the possible protein variant space in comparison to other systems. Previously, if one generated, for example, 10^{13} chimeric molecules during a reaction, it would be extremely difficult to test such a high number of chimeric variants for a particular activity. Moreover, a significant portion of the progeny population would have a very high number of crossover events which resulted in proteins that were less likely to have increased levels of a particular activity. By using these methods, the population of chimerics molecules can be enriched for those variants that have a particular number of crossover events. Thus, although one can still generate 10^{13} chimeric molecules during a reaction, each of the molecules chosen for further analysis most likely has, for example, only three crossover events. Because the resulting progeny population can be skewed to have a predetermined number of crossover events, the boundaries on the functional variety between the chimeric molecules is reduced. This provides a more manageable number of variables when calculating which oligonucleotide from the original parental polynucleotides might be responsible for affecting a particular trait.

One method for creating a chimeric progeny polynucleotide sequence is to create oligonucleotides corresponding to fragments or portions of each parental sequence. Each oligonucleotide preferably includes a unique region of overlap so that mixing the oligonucleotides together results in a new variant that has each oligonucleotide fragment assembled in the correct order. Additional information can also be found, e.g., in USSN 09/332,835; U.S. Patent No. 6,361,974.

The number of oligonucleotides generated for each parental variant bears a relationship to the total number of resulting crossovers in the chimeric molecule that is ultimately created. For example, three parental nucleotide sequence variants might be provided to undergo a ligation reaction in order to find a chimeric variant having, for example, greater activity at high temperature. As one example, a set of 50 oligonucleotide sequences can be generated corresponding to each portions of each parental variant. Accordingly, during the ligation reassembly process there could be up to 50 crossover events within each of the chimeric sequences. The probability that each of the generated chimeric polynucleotides will contain oligonucleotides from each parental variant in alternating order is very low. If each oligonucleotide fragment is present in the ligation reaction in the same molar quantity it is likely that in some positions oligonucleotides from the same parental polynucleotide will ligate next to one another

and thus not result in a crossover event. If the concentration of each oligonucleotide from each parent is kept constant during any ligation step in this example, there is a 1/3 chance (assuming 3 parents) that an oligonucleotide from the same parental variant will ligate within the chimeric sequence and produce no crossover.

Accordingly, a probability density function (PDF) can be determined to predict the population of crossover events that are likely to occur during each step in a ligation reaction given a set number of parental variants, a number of oligonucleotides corresponding to each variant, and the concentrations of each variant during each step in the ligation reaction. The statistics and mathematics behind determining the PDF is described below. By utilizing these methods, one can calculate such a probability density function, and thus enrich the chimeric progeny population for a predetermined number of crossover events resulting from a particular ligation reaction. Moreover, a target number of crossover events can be predetermined, and the system then programmed to calculate the starting quantities of each parental oligonucleotide during each step in the ligation reaction to result in a probability density function that centers on the predetermined number of crossover events. These methods are directed to the use of repeated cycles of reductive reassortment, recombination and selection that allow for the directed molecular evolution of a nucleic acid encoding a polypeptide through recombination. This system allows generation of a large population of evolved chimeric sequences, wherein the generated population is significantly enriched for sequences that have a predetermined number of crossover events. A crossover event is a point in a chimeric sequence where a shift in sequence occurs from one parental variant to another parental variant. Such a point is normally at the juncture of where oligonucleotides from two parents are ligated together to form a single sequence. The method allows calculation of the correct concentrations of oligonucleotide sequences so that the final chimeric population of sequences is enriched for the chosen number of crossover events. This provides more control over choosing chimeric variants having a predetermined number of crossover events.

In addition, these methods provide a convenient means for exploring a tremendous amount of the possible protein variant space in comparison to other systems. By using the methods described herein, the population of chimerics molecules can be enriched for those variants that have a particular number of crossover events. Thus, although one can still generate 10^{13} chimeric molecules during a reaction, each of the

molecules chosen for further analysis most likely has, for example, only three crossover events. Because the resulting progeny population can be skewed to have a predetermined number of crossover events, the boundaries on the functional variety between the chimeric molecules is reduced. This provides a more manageable number of variables when calculating which oligonucleotide from the original parental polynucleotides might be responsible for affecting a particular trait.

In one aspect, the method creates a chimeric progeny polynucleotide sequence by creating oligonucleotides corresponding to fragments or portions of each parental sequence. Each oligonucleotide preferably includes a unique region of overlap so that mixing the oligonucleotides together results in a new variant that has each oligonucleotide fragment assembled in the correct order. See also USSN 09/332,835.

Determining Crossover Events

Aspects of the invention include a system and software that receive a desired crossover probability density function (PDF), the number of parent genes to be reassembled, and the number of fragments in the reassembly as inputs. The output of this program is a "fragment PDF" that can be used to determine a recipe for producing reassembled genes, and the estimated crossover PDF of those genes. The processing described herein is preferably performed in MATLAB™ (The Mathworks, Natick, Massachusetts) a programming language and development environment for technical computing.

Iterative Processes

In practicing the invention, these processes can be iteratively repeated. For example, a nucleic acid (or, the nucleic acid) responsible for an altered or new amylase phenotype is identified, re-isolated, again modified, re-tested for activity. This process can be iteratively repeated until a desired phenotype is engineered. For example, an entire biochemical anabolic or catabolic pathway can be engineered into a cell, including, e.g., starch hydrolysis activity.

Similarly, if it is determined that a particular oligonucleotide has no affect at all on the desired trait (e.g., a new amylase phenotype), it can be removed as a variable by synthesizing larger parental oligonucleotides that include the sequence to be removed. Since incorporating the sequence within a larger sequence prevents any crossover events, there will no longer be any variation of this sequence in the progeny polynucleotides.

This iterative practice of determining which oligonucleotides are most related to the desired trait, and which are unrelated, allows more efficient exploration all of the possible protein variants that might be provide a particular trait or activity.

In vivo shuffling

In vivo shuffling of molecules is use in methods of the invention that provide variants of polypeptides of the invention, e.g., antibodies, amylases, and the like. *In vivo* shuffling can be performed utilizing the natural property of cells to recombine multimers. While recombination *in vivo* has provided the major natural route to molecular diversity, genetic recombination remains a relatively complex process that involves 1) the recognition of homologies; 2) strand cleavage, strand invasion, and metabolic steps leading to the production of recombinant chiasma; and finally 3) the resolution of chiasma into discrete recombined molecules. The formation of the chiasma requires the recognition of homologous sequences.

In one aspect, the invention provides a method for producing a hybrid polynucleotide from at least a first polynucleotide (e.g., an amylase of the invention) and a second polynucleotide (e.g., an enzyme, such as an amylase of the invention or any other amylase, or, a tag or an epitope). The invention can be used to produce a hybrid polynucleotide by introducing at least a first polynucleotide and a second polynucleotide which share at least one region of partial sequence homology into a suitable host cell. The regions of partial sequence homology promote processes which result in sequence reorganization producing a hybrid polynucleotide. The term "hybrid polynucleotide", as used herein, is any nucleotide sequence which results from the method of the present invention and contains sequence from at least two original polynucleotide sequences. Such hybrid polynucleotides can result from intermolecular recombination events which promote sequence integration between DNA molecules. In addition, such hybrid polynucleotides can result from intramolecular reductive reassortment processes which utilize repeated sequences to alter a nucleotide sequence within a DNA molecule.

Producing sequence variants

The invention also provides additional methods for making sequence variants of the nucleic acid (e.g., amylase) sequences of the invention. The invention also provides additional methods for isolating amylases using the nucleic acids and polypeptides of the invention. In one aspect, the invention provides for variants of an

amylase coding sequence (e.g., a gene, cDNA or message) of the invention, which can be altered by any means, including, e.g., random or stochastic methods, or, non-stochastic, or "directed evolution," methods, as described above.

The isolated variants may be naturally occurring. Variant can also be created *in vitro*. Variants may be created using genetic engineering techniques such as site directed mutagenesis, random chemical mutagenesis, Exonuclease III deletion procedures, and standard cloning techniques. Alternatively, such variants, fragments, analogs, or derivatives may be created using chemical synthesis or modification procedures. Other methods of making variants are also familiar to those skilled in the art. These include procedures in which nucleic acid sequences obtained from natural isolates are modified to generate nucleic acids which encode polypeptides having characteristics which enhance their value in industrial or laboratory applications. In such procedures, a large number of variant sequences having one or more nucleotide differences with respect to the sequence obtained from the natural isolate are generated and characterized. These nucleotide differences can result in amino acid changes with respect to the polypeptides encoded by the nucleic acids from the natural isolates.

For example, variants may be created using error prone PCR. In error prone PCR, PCR is performed under conditions where the copying fidelity of the DNA polymerase is low, such that a high rate of point mutations is obtained along the entire length of the PCR product. Error prone PCR is described, e.g., in Leung, D.W., et al., Technique, 1:11-15, 1989) and Caldwell, R. C. & Joyce G.F., PCR Methods Applic., 2:28-33, 1992. Briefly, in such procedures, nucleic acids to be mutagenized are mixed with PCR primers, reaction buffer, $MgCl_2$, $MnCl_2$, Taq polymerase and an appropriate concentration of dNTPs for achieving a high rate of point mutation along the entire length of the PCR product. For example, the reaction may be performed using 20 fmoles of nucleic acid to be mutagenized, 30 pmole of each PCR primer, a reaction buffer comprising 50mM KCl, 10mM Tris HCl (pH 8.3) and 0.01% gelatin, 7mM MgCl2, 0.5mM $MnCl_2$, 5 units of Taq polymerase, 0.2mM dGTP, 0.2mM dATP, 1mM dCTP, and 1mM dTTP. PCR may be performed for 30 cycles of 94°C for 1 min, 45°C for 1 min, and 72°C for 1 min. However, it will be appreciated that these parameters may be varied as appropriate. The mutagenized nucleic acids are cloned into an appropriate vector and the activities of the polypeptides encoded by the mutagenized nucleic acids is evaluated.

Variants may also be created using oligonucleotide directed mutagenesis to generate site-specific mutations in any cloned DNA of interest. Oligonucleotide mutagenesis is described, e.g., in Reidhaar-Olson (1988) Science 241:53-57. Briefly, in such procedures a plurality of double stranded oligonucleotides bearing one or more mutations to be introduced into the cloned DNA are synthesized and inserted into the cloned DNA to be mutagenized. Clones containing the mutagenized DNA are recovered and the activities of the polypeptides they encode are assessed.

Another method for generating variants is assembly PCR. Assembly PCR involves the assembly of a PCR product from a mixture of small DNA fragments. A large number of different PCR reactions occur in parallel in the same vial, with the products of one reaction priming the products of another reaction. Assembly PCR is described in, e.g., U.S. Patent No. 5,965,408.

Still another method of generating variants is sexual PCR mutagenesis. In sexual PCR mutagenesis, forced homologous recombination occurs between DNA molecules of different but highly related DNA sequence in vitro, as a result of random fragmentation of the DNA molecule based on sequence homology, followed by fixation of the crossover by primer extension in a PCR reaction. Sexual PCR mutagenesis is described, e.g., in Stemmer (1994) Proc. Natl. Acad. Sci. USA 91:10747-10751. Briefly, in such procedures a plurality of nucleic acids to be recombined are digested with DNase to generate fragments having an average size of 50-200 nucleotides. Fragments of the desired average size are purified and resuspended in a PCR mixture. PCR is conducted under conditions which facilitate recombination between the nucleic acid fragments. For example, PCR may be performed by resuspending the purified fragments at a concentration of 10-30ng/:l in a solution of 0.2mM of each dNTP, 2.2mM $MgCl_2$, 50mM KCL, 10mM Tris HCl, pH 9.0, and 0.1% Triton X-100. 2.5 units of Taq polymerase per 100:l of reaction mixture is added and PCR is performed using the following regime: 94°C for 60 seconds, 94°C for 30 seconds, 50-55°C for 30 seconds, 72°C for 30 seconds (30-45 times) and 72°C for 5 minutes. However, it will be appreciated that these parameters may be varied as appropriate. In some aspects, oligonucleotides may be included in the PCR reactions. In other aspects, the Klenow fragment of DNA polymerase I may be used in a first set of PCR reactions and Taq polymerase may be used in a subsequent set of PCR reactions. Recombinant sequences are isolated and the activities of the polypeptides they encode are assessed.

Variants may also be created by *in vivo* mutagenesis. In some aspects, random mutations in a sequence of interest are generated by propagating the sequence of interest in a bacterial strain, such as an *E. coli* strain, which carries mutations in one or more of the DNA repair pathways. Such "mutator" strains have a higher random mutation rate than that of a wild-type parent. Propagating the DNA in one of these strains will eventually generate random mutations within the DNA. Mutator strains suitable for use for in vivo mutagenesis are described, e.g., in PCT Publication No. WO 91/16427.

Variants may also be generated using cassette mutagenesis. In cassette mutagenesis a small region of a double stranded DNA molecule is replaced with a synthetic oligonucleotide "cassette" that differs from the native sequence. The oligonucleotide often contains completely and/or partially randomized native sequence.

Recursive ensemble mutagenesis may also be used to generate variants. Recursive ensemble mutagenesis is an algorithm for protein engineering (protein mutagenesis) developed to produce diverse populations of phenotypically related mutants whose members differ in amino acid sequence. This method uses a feedback mechanism to control successive rounds of combinatorial cassette mutagenesis. Recursive ensemble mutagenesis is described, e.g., in Arkin (1992) Proc. Natl. Acad. Sci. USA 89:7811-7815.

In some aspects, variants are created using exponential ensemble mutagenesis. Exponential ensemble mutagenesis is a process for generating combinatorial libraries with a high percentage of unique and functional mutants, wherein small groups of residues are randomized in parallel to identify, at each altered position, amino acids which lead to functional proteins. Exponential ensemble mutagenesis is described, e.g., in Delegrave (1993) Biotechnology Res. 11:1548-1552. Random and site-directed mutagenesis are described, e.g., in Arnold (1993) Current Opinion in Biotechnology 4:450-455.

In some aspects, the variants are created using shuffling procedures wherein portions of a plurality of nucleic acids which encode distinct polypeptides are fused together to create chimeric nucleic acid sequences which encode chimeric polypeptides as described in, e.g., U.S. Patent Nos. 5,965,408; 5,939,250 (see also discussion, above).

The invention also provides variants of polypeptides of the invention (e.g., amylases) comprising sequences in which one or more of the amino acid residues (e.g., of an exemplary polypeptide of the invention) are substituted with a conserved or non-

conserved amino acid residue (e.g., a conserved amino acid residue) and such substituted amino acid residue may or may not be one encoded by the genetic code. Conservative substitutions are those that substitute a given amino acid in a polypeptide by another amino acid of like characteristics. Thus, polypeptides of the invention include those with conservative substitutions of sequences of the invention, e.g., the exemplary polypeptides of the invention, including but not limited to the following replacements: replacements of an aliphatic amino acid such as Alanine, Valine, Leucine and Isoleucine with another aliphatic amino acid; replacement of a Serine with a Threonine or vice versa; replacement of an acidic residue such as Aspartic acid and Glutamic acid with another acidic residue; replacement of a residue bearing an amide group, such as Asparagine and Glutamine, with another residue bearing an amide group; exchange of a basic residue such as Lysine and Arginine with another basic residue; and replacement of an aromatic residue such as Phenylalanine, Tyrosine with another aromatic residue. Other variants are those in which one or more of the amino acid residues of the polypeptides of the invention includes a substituent group.

Other variants within the scope of the invention are those in which the polypeptide is associated with another compound, such as a compound to increase the half-life of the polypeptide, for example, polyethylene glycol.

Additional variants within the scope of the invention are those in which additional amino acids are fused to the polypeptide, such as a leader sequence, a secretory sequence, a proprotein sequence or a sequence which facilitates purification, enrichment, or stabilization of the polypeptide.

In some aspects, the variants, fragments, derivatives and analogs of the polypeptides of the invention retain the same biological function or activity as the exemplary polypeptides, e.g., amylase activity, as described herein. In other aspects, the variant, fragment, derivative, or analog includes a proprotein, such that the variant, fragment, derivative, or analog can be activated by cleavage of the proprotein portion to produce an active polypeptide.

Optimizing codons to achieve high levels of protein expression in host cells

The invention provides methods for modifying amylase-encoding nucleic acids to modify codon usage. In one aspect, the invention provides methods for modifying codons in a nucleic acid encoding an amylase to increase or decrease its

expression in a host cell. The invention also provides nucleic acids encoding an amylase modified to increase its expression in a host cell, amylase so modified, and methods of making the modified amylases. The method comprises identifying a "non-preferred" or a "less preferred" codon in amylase-encoding nucleic acid and replacing one or more of these non-preferred or less preferred codons with a "preferred codon" encoding the same amino acid as the replaced codon and at least one non-preferred or less preferred codon in the nucleic acid has been replaced by a preferred codon encoding the same amino acid. A preferred codon is a codon over-represented in coding sequences in genes in the host cell and a non-preferred or less preferred codon is a codon under-represented in coding sequences in genes in the host cell.

Host cells for expressing the nucleic acids, expression cassettes and vectors of the invention include bacteria, yeast, fungi, plant cells, insect cells and mammalian cells. Thus, the invention provides methods for optimizing codon usage in all of these cells, codon-altered nucleic acids and polypeptides made by the codon-altered nucleic acids. Exemplary host cells include gram negative bacteria, such as *Escherichia coli*; gram positive bacteria, such as *Bacillus cereus*, *Streptomyces*, *Lactobacillus gasseri*, *Lactococcus lactis*, *Lactococcus cremoris*, *Bacillus subtilis*. Exemplary host cells also include eukaryotic organisms, e.g., various yeast, such as *Saccharomyces* sp., including *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Pichia pastoris*, and *Kluyveromyces lactis*, *Hansenula polymorpha*, *Aspergillus niger*, and mammalian cells and cell lines and insect cells and cell lines. Thus, the invention also includes nucleic acids and polypeptides optimized for expression in these organisms and species.

For example, the codons of a nucleic acid encoding an amylase isolated from a bacterial cell are modified such that the nucleic acid is optimally expressed in a bacterial cell different from the bacteria from which the amylase was derived, a yeast, a fungi, a plant cell, an insect cell or a mammalian cell. Methods for optimizing codons are well known in the art, see, e.g., U.S. Patent No. 5,795,737; Baca (2000) Int. J. Parasitol. 30:113-118; Hale (1998) Protein Expr. Purif. 12:185-188; Narum (2001) Infect. Immun. 69:7250-7253. See also Narum (2001) Infect. Immun. 69:7250-7253, describing optimizing codons in mouse systems; Outchkourov (2002) Protein Expr. Purif. 24:18-24, describing optimizing codons in yeast; Feng (2000) Biochemistry 39:15399-15409, describing optimizing codons in *E. coli*; Humphreys (2000) Protein Expr. Purif. 20:252-264, describing optimizing codon usage that affects secretion in *E. coli*.

Transgenic non-human animals

The invention provides transgenic non-human animals comprising a nucleic acid, a polypeptide (e.g., an amylase), an expression cassette or vector or a transfected or transformed cell of the invention. The invention also provides methods of making and using these transgenic non-human animals.

The transgenic non-human animals can be, e.g., goats, rabbits, sheep, pigs, cows, rats and mice, comprising the nucleic acids of the invention. These animals can be used, e.g., as *in vivo* models to study amylase activity, or, as models to screen for agents that change the amylase activity *in vivo*. The coding sequences for the polypeptides to be expressed in the transgenic non-human animals can be designed to be constitutive, or, under the control of tissue-specific, developmental-specific or inducible transcriptional regulatory factors. Transgenic non-human animals can be designed and generated using any method known in the art; see, e.g., U.S. Patent Nos. 6,211,428; 6,187,992; 6,156,952; 6,118,044; 6,111,166; 6,107,541; 5,959,171; 5,922,854; 5,892,070; 5,880,327; 5,891,698; 5,639,940; 5,573,933; 5,387,742; 5,087,571, describing making and using transformed cells and eggs and transgenic mice, rats, rabbits, sheep, pigs and cows. See also, e.g., Pollock (1999) J. Immunol. Methods 231:147-157, describing the production of recombinant proteins in the milk of transgenic dairy animals; Baguisi (1999) Nat. Biotechnol. 17:456-461, demonstrating the production of transgenic goats. U.S. Patent No. 6,211,428, describes making and using transgenic non-human mammals which express in their brains a nucleic acid construct comprising a DNA sequence. U.S. Patent No. 5,387,742, describes injecting cloned recombinant or synthetic DNA sequences into fertilized mouse eggs, implanting the injected eggs in pseudo-pregnant females, and growing to term transgenic mice whose cells express proteins related to the pathology of Alzheimer's disease. U.S. Patent No. 6,187,992, describes making and using a transgenic mouse whose genome comprises a disruption of the gene encoding amyloid precursor protein (APP).

"Knockout animals" can also be used to practice the methods of the invention. For example, in one aspect, the transgenic or modified animals of the invention comprise a "knockout animal," e.g., a "knockout mouse," engineered not to express an endogenous gene, which is replaced with a gene expressing an amylase of the invention, or, a fusion protein comprising an amylase of the invention.

Transgenic Plants and Seeds

The invention provides transgenic plants and seeds comprising a nucleic acid, a polypeptide (e.g., an amylase, such as an alpha amylase), an expression cassette or vector or a transfected or transformed cell of the invention. The invention also provides plant products, e.g., oils, seeds, leaves, extracts and the like, comprising a nucleic acid and/or a polypeptide (e.g., an amylase, such as an alpha amylase) of the invention. The transgenic plant can be dicotyledonous (a dicot) or monocotyledonous (a monocot). The invention also provides methods of making and using these transgenic plants and seeds. The transgenic plant or plant cell expressing a polypeptide of the present invention may be constructed in accordance with any method known in the art. See, for example, U.S. Patent No. 6,309,872.

Nucleic acids and expression constructs of the invention can be introduced into a plant cell by any means. For example, nucleic acids or expression constructs can be introduced into the genome of a desired plant host, or, the nucleic acids or expression constructs can be episomes. Introduction into the genome of a desired plant can be such that the host's a-amylase production is regulated by endogenous transcriptional or translational control elements. The invention also provides "knockout plants" where insertion of gene sequence by, e.g., homologous recombination, has disrupted the expression of the endogenous gene. Means to generate "knockout" plants are well-known in the art, see, e.g., Strepp (1998) Proc Natl. Acad. Sci. USA 95:4368-4373; Miao (1995) Plant J 7:359-365. See discussion on transgenic plants, below.

The nucleic acids of the invention can be used to confer desired traits on essentially any plant, e.g., on starch-producing plants, such as potato, wheat, rice, barley, and the like. Nucleic acids of the invention can be used to manipulate metabolic pathways of a plant in order to optimize or alter host's expression of a-amylase. The can change the ratio of starch/sugar conversion in a plant. This can facilitate industrial processing of a plant. Alternatively, alpha-amylases of the invention can be used in production of a transgenic plant to produce a compound not naturally produced by that plant. This can lower production costs or create a novel product.

In one aspect, the first step in production of a transgenic plant involves making an expression construct for expression in a plant cell. These techniques are well known in the art. They can include selecting and cloning a promoter, a coding sequence for facilitating efficient binding of ribosomes to mRNA and selecting the appropriate

gene terminator sequences. One exemplary constitutive promoter is CaMV35S, from the cauliflower mosaic virus, which generally results in a high degree of expression in plants. Other promoters are more specific and respond to cues in the plant's internal or external environment. An exemplary light-inducible promoter is the promoter from the cab gene, encoding the major chlorophyll a/b binding protein.

In one aspect, the nucleic acid is modified to achieve greater expression in a plant cell. For example, a sequence of the invention is likely to have a higher percentage of A-T nucleotide pairs compared to that seen in a plant, some of which prefer G-C nucleotide pairs. Therefore, A-T nucleotides in the coding sequence can be substituted with G-C nucleotides without significantly changing the amino acid sequence to enhance production of the gene product in plant cells.

Selectable marker gene can be added to the gene construct in order to identify plant cells or tissues that have successfully integrated the transgene. This may be necessary because achieving incorporation and expression of genes in plant cells is a rare event, occurring in just a few percent of the targeted tissues or cells. Selectable marker genes encode proteins that provide resistance to agents that are normally toxic to plants, such as antibiotics or herbicides. Only plant cells that have integrated the selectable marker gene will survive when grown on a medium containing the appropriate antibiotic or herbicide. As for other inserted genes, marker genes also require promoter and termination sequences for proper function.

In one aspect, making transgenic plants or seeds comprises incorporating sequences of the invention and, optionally, marker genes into a target expression construct (e.g., a plasmid), along with positioning of the promoter and the terminator sequences. This can involve transferring the modified gene into the plant through a suitable method. For example, a construct may be introduced directly into the genomic DNA of the plant cell using techniques such as electroporation and microinjection of plant cell protoplasts, or the constructs can be introduced directly to plant tissue using ballistic methods, such as DNA particle bombardment. For example, see, e.g., Christou (1997) Plant Mol. Biol. 35:197-203; Pawlowski (1996) Mol. Biotechnol. 6:17-30; Klein (1987) Nature 327:70-73; Takumi (1997) Genes Genet. Syst. 72:63-69, discussing use of particle bombardment to introduce transgenes into wheat; and Adam (1997) supra, for use of particle bombardment to introduce YACs into plant cells. For example, Rinehart (1997) supra, used particle bombardment to generate transgenic cotton plants. Apparatus

for accelerating particles is described U.S. Pat. No. 5,015,580; and, the commercially available BioRad (Biolistics) PDS-2000 particle acceleration instrument; see also, John, U.S. Patent No. 5,608,148; and Ellis, U.S. Patent No. 5, 681,730, describing particle-mediated transformation of gymnosperms.

In one aspect, protoplasts can be immobilized and injected with a nucleic acids, e.g., an expression construct. Although plant regeneration from protoplasts is not easy with cereals, plant regeneration is possible in legumes using somatic embryogenesis from protoplast derived callus. Organized tissues can be transformed with naked DNA using gene gun technique, where DNA is coated on tungsten microprojectiles, shot 1/100th the size of cells, which carry the DNA deep into cells and organelles. Transformed tissue is then induced to regenerate, usually by somatic embryogenesis. This technique has been successful in several cereal species including maize and rice.

Nucleic acids, e.g., expression constructs, can also be introduced in to plant cells using recombinant viruses. Plant cells can be transformed using viral vectors, such as, e.g., tobacco mosaic virus derived vectors (Rouwendal (1997) Plant Mol. Biol. 33:989-999), see Porta (1996) "Use of viral replicons for the expression of genes in plants," Mol. Biotechnol. 5:209-221.

Alternatively, nucleic acids, e.g., an expression construct, can be combined with suitable T-DNA flanking regions and introduced into a conventional *Agrobacterium tumefaciens* host vector. The virulence functions of the *Agrobacterium tumefaciens* host will direct the insertion of the construct and adjacent marker into the plant cell DNA when the cell is infected by the bacteria. *Agrobacterium tumefaciens*-mediated transformation techniques, including disarming and use of binary vectors, are well described in the scientific literature. See, e.g., Horsch (1984) *Science* 233:496-498; Fraley (1983) *Proc. Natl. Acad. Sci. USA* 80:4803 (1983); *Gene Transfer to Plants*, Potrykus, ed. (Springer-Verlag, Berlin 1995). The DNA in an A. *tumefaciens* cell is contained in the bacterial chromosome as well as in another structure known as a Ti (tumor-inducing) plasmid. The Ti plasmid contains a stretch of DNA termed T-DNA (~20 kb long) that is transferred to the plant cell in the infection process and a series of vir (virulence) genes that direct the infection process. *A. tumefaciens* can only infect a plant through wounds: when a plant root or stem is wounded it gives off certain chemical signals, in response to which, the vir genes of *A. tumefaciens* become activated and direct a series of events necessary for the transfer of the T-DNA from the Ti plasmid to the

plant's chromosome. The T-DNA then enters the plant cell through the wound. One speculation is that the T-DNA waits until the plant DNA is being replicated or transcribed, then inserts itself into the exposed plant DNA. In order to use *A. tumefaciens* as a transgene vector, the tumor-inducing section of T-DNA have to be removed, while retaining the T-DNA border regions and the vir genes. The transgene is then inserted between the T-DNA border regions, where it is transferred to the plant cell and becomes integrated into the plant's chromosomes.

The invention provides for the transformation of monocotyledonous plants using the nucleic acids of the invention, including important cereals, see Hiei (1997) Plant Mol. Biol. 35:205-218. See also, *e.g.*, Horsch, Science (1984) 233:496; Fraley (1983) Proc. Natl Acad. Sci USA 80:4803; Thykjaer (1997) supra; Park (1996) Plant Mol. Biol. 32:1135-1148, discussing T-DNA integration into genomic DNA. See also D'Halluin, U.S. Patent No. 5,712,135, describing a process for the stable integration of a DNA comprising a gene that is functional in a cell of a cereal, or other monocotyledonous plant.

In one aspect, the third step can involve selection and regeneration of whole plants capable of transmitting the incorporated target gene to the next generation. Such regeneration techniques rely on manipulation of certain phytohormones in a tissue culture growth medium, typically relying on a biocide and/or herbicide marker that has been introduced together with the desired nucleotide sequences. Plant regeneration from cultured protoplasts is described in Evans et al., *Protoplasts Isolation and Culture, Handbook of Plant Cell Culture*, pp. 124-176, MacMillilan Publishing Company, New York, 1983; and Binding, *Regeneration of Plants, Plant Protoplasts*, pp. 21-73, CRC Press, Boca Raton, 1985. Regeneration can also be obtained from plant callus, explants, organs, or parts thereof. Such regeneration techniques are described generally in Klee (1987) Ann. Rev. of Plant Phys. 38:467-486. To obtain whole plants from transgenic tissues such as immature embryos, they can be grown under controlled environmental conditions in a series of media containing nutrients and hormones, a process known as tissue culture. Once whole plants are generated and produce seed, evaluation of the progeny begins.

After the expression cassette is stably incorporated in transgenic plants, it can be introduced into other plants by sexual crossing. Any of a number of standard breeding techniques can be used, depending upon the species to be crossed. Since

transgenic expression of the nucleic acids of the invention leads to phenotypic changes, plants comprising the recombinant nucleic acids of the invention can be sexually crossed with a second plant to obtain a final product. Thus, the seed of the invention can be derived from a cross between two transgenic plants of the invention, or a cross between a plant of the invention and another plant. The desired effects (e.g., expression of the polypeptides of the invention to produce a plant in which flowering behavior is altered) can be enhanced when both parental plants express the polypeptides (e.g., an amylase, such as an alpha amylase) of the invention. The desired effects can be passed to future plant generations by standard propagation means.

The nucleic acids and polypeptides of the invention are expressed in or inserted in any plant or seed. Transgenic plants of the invention can be dicotyledonous or monocotyledonous. Examples of monocot transgenic plants of the invention are grasses, such as meadow grass (blue grass, *Poa*), forage grass such as festuca, lolium, temperate grass, such as *Agrostis*, and cereals, e.g., wheat, oats, rye, barley, rice, sorghum, and maize (corn). Examples of dicot transgenic plants of the invention are tobacco, legumes, such as lupins, potato, sugar beet, pea, bean and soybean, and cruciferous plants (family *Brassicaceae*), such as cauliflower, rape seed, and the closely related model organism *Arabidopsis thaliana*. Thus, the transgenic plants and seeds of the invention include a broad range of plants, including, but not limited to, species from the genera *Anacardium, Arachis, Asparagus, Atropa, Avena, Brassica, Citrus, Citrullus, Capsicum, Carthamus, Cocos, Coffea, Cucumis, Cucurbita, Daucus, Elaeis, Fragaria, Glycine, Gossypium, Helianthus, Heterocallis, Hordeum, Hyoscyamus, Lactuca, Linum, Lolium, Lupinus, Lycopersicon, Malus, Manihot, Majorana, Medicago, Nicotiana, Olea, Oryza, Panieum, Pannisetum, Persea, Phaseolus, Pistachia, Pisum, Pyrus, Prunus, Raphanus, Ricinus, Secale, Senecio, Sinapis, Solanum, Sorghum, Theobromus, Trigonella, Triticum, Vicia, Vitis, Vigna*, and *Zea*.

In alternative embodiments, the nucleic acids of the invention are expressed in plants which contain fiber cells, including, *e.g.*, cotton, silk *cotton* tree (Kapok, Ceiba pentandra), desert willow, creosote bush, winterfat, balsa, ramie, kenaf, hemp, roselle, jute, sisal abaca and flax. In alternative embodiments, the transgenic plants of the invention can be members of the genus *Gossypium*, including members of any *Gossypium* species, such as *G. arboreum*;. *G. herbaceum*, *G. barbadense*, and *G. hirsutum*.

The invention also provides for transgenic plants to be used for producing large amounts of the polypeptides (e.g., an amylase, such as an alpha amylase) of the invention. For example, see Palmgren (1997) Trends Genet. 13:348; Chong (1997) Transgenic Res. 6:289-296 (producing human milk protein beta-casein in transgenic potato plants using an auxin-inducible, bidirectional mannopine synthase (mas1',2') promoter with *Agrobacterium tumefaciens*-mediated leaf disc transformation methods).

Using known procedures, one of skill can screen for plants of the invention by detecting the increase or decrease of transgene mRNA or protein in transgenic plants. Means for detecting and quantitation of mRNAs or proteins are well known in the art.

Polypeptides and peptides

In one aspect, the invention provides isolated or recombinant polypeptides having a sequence identity (e.g., at least about 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more, or complete (100%) sequence identity) to an exemplary sequence of the invention, e.g., SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:20, SEQ ID NO:22, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32, SEQ ID NO:34, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:44, SEQ ID NO:46, SEQ ID NO:48, SEQ ID NO:50, SEQ ID NO:52, SEQ ID NO:54, SEQ ID NO:56, SEQ ID NO:58, SEQ ID NO:60, SEQ ID NO:62, SEQ ID NO:64, SEQ ID NO:66, SEQ ID NO:68, SEQ ID NO:70, SEQ ID NO:72, SEQ ID NO:74, SEQ ID NO:76, SEQ ID NO:78, SEQ ID NO:80, SEQ ID NO:82, SEQ ID NO:84, SEQ ID NO:86, SEQ ID NO:88, SEQ ID NO:90, SEQ ID NO:92, SEQ ID NO:94, SEQ ID NO:96, SEQ ID NO:98, SEQ ID NO:100, SEQ ID NO:102, SEQ ID NO:104, SEQ ID NO:106, SEQ ID NO:108, SEQ ID NO:110, SEQ ID NO:112, SEQ ID NO:114, SEQ ID NO:116, SEQ ID NO:118, SEQ ID NO:120, SEQ ID NO:122, SEQ ID NO:124, SEQ ID NO:126, SEQ ID NO:128, SEQ ID NO:130, SEQ ID NO:132, SEQ ID NO:134, SEQ ID NO:136, SEQ ID NO:138, SEQ ID NO:140, SEQ ID NO:142, SEQ ID NO:144, SEQ ID NO:146, SEQ ID NO:148, SEQ ID NO:150, SEQ ID NO:152, SEQ ID NO:154, SEQ ID NO:156, SEQ ID NO:158, SEQ ID NO:160, SEQ ID NO:162, SEQ ID NO:164, SEQ ID

NO:166, SEQ ID NO:168, SEQ ID NO:190, SEQ ID NO:192, SEQ ID NO:194, SEQ ID NO:204, SEQ ID NO:206, SEQ ID NO:208, SEQ ID NO:210, SEQ ID NO:212, SEQ ID NO:323, SEQ ID NO:325, SEQ ID NO:327, SEQ ID NO:329, SEQ ID NO:331, SEQ ID NO:333, SEQ ID NO:335, SEQ ID NO:337, SEQ ID NO:339, SEQ ID NO:341, SEQ ID NO:343, SEQ ID NO:345, SEQ ID NO:347, SEQ ID NO:349, SEQ ID NO:351, SEQ ID NO:353, SEQ ID NO:355, SEQ ID NO:357, SEQ ID NO:359, SEQ ID NO:361, SEQ ID NO:363, SEQ ID NO:365, SEQ ID NO:367, SEQ ID NO:369, SEQ ID NO:371, SEQ ID NO:373, SEQ ID NO:375, SEQ ID NO:377, SEQ ID NO:379, SEQ ID NO:381, SEQ ID NO:383, SEQ ID NO:385, SEQ ID NO:387, SEQ ID NO:389, SEQ ID NO:391, SEQ ID NO:393, SEQ ID NO:395, SEQ ID NO:397, SEQ ID NO:399, SEQ ID NO:401, SEQ ID NO:403, SEQ ID NO:405, SEQ ID NO:407, SEQ ID NO:409, SEQ ID NO:411, SEQ ID NO:413, SEQ ID NO:415, SEQ ID NO:417, SEQ ID NO:419, SEQ ID NO:421, SEQ ID NO:423, SEQ ID NO:425, SEQ ID NO:427, SEQ ID NO:429, SEQ ID NO:431, SEQ ID NO:433, SEQ ID NO:435, SEQ ID NO:437, SEQ ID NO:439, SEQ ID NO:441, SEQ ID NO:443, SEQ ID NO:445, SEQ ID NO:447, SEQ ID NO:449, SEQ ID NO:451, SEQ ID NO:453, SEQ ID NO:455, SEQ ID NO:457, SEQ ID NO:459, SEQ ID NO:461, SEQ ID NO:461, SEQ ID NO:463, SEQ ID NO:464, SEQ ID NO:466, SEQ ID NO:468, SEQ ID NO:469, SEQ ID NO:470, SEQ ID NO:471, SEQ ID NO:472, SEQ ID NO:474, SEQ ID NO:476, SEQ ID NO:477, SEQ ID NO:479, SEQ ID NO:481, SEQ ID NO:482, SEQ ID NO:483, SEQ ID NO:485, SEQ ID NO:487, SEQ ID NO:488, SEQ ID NO:489, SEQ ID NO:490, SEQ ID NO:491, SEQ ID NO:493, SEQ ID NO:495, SEQ ID NO:496, SEQ ID NO:497, SEQ ID NO:499, SEQ ID NO:501, SEQ ID NO:502, SEQ ID NO:503, SEQ ID NO:504, SEQ ID NO:505, SEQ ID NO:506, SEQ ID NO:507, SEQ ID NO:508, SEQ ID NO:510, SEQ ID NO:512, SEQ ID NO:513, SEQ ID NO:514, SEQ ID NO:516, SEQ ID NO:518, SEQ ID NO:518, SEQ ID NO:520, SEQ ID NO:521, SEQ ID NO:523, SEQ ID NO:525, SEQ ID NO:526, SEQ ID NO:528, SEQ ID NO:530, SEQ ID NO:531, SEQ ID NO:533, SEQ ID NO:535, SEQ ID NO:536, SEQ ID NO:537, SEQ ID NO:538, SEQ ID NO:540, SEQ ID NO:542, SEQ ID NO:543, SEQ ID NO:545, SEQ ID NO:547, SEQ ID NO:548, SEQ ID NO:549, SEQ ID NO:550, SEQ ID NO:551, SEQ ID NO:553, SEQ ID NO:555, SEQ ID NO:556, SEQ ID NO:557, SEQ ID NO:559, SEQ ID NO:561, SEQ ID NO:562, SEQ ID NO:563, SEQ ID NO:564, SEQ ID NO:566, SEQ ID NO:568, SEQ ID NO:570, SEQ ID NO:572, SEQ ID NO:574, SEQ ID NO:576, SEQ ID NO:578, SEQ ID NO:580, SEQ ID NO:582, SEQ ID NO:584, SEQ ID NO:586, SEQ ID NO:588, SEQ ID

NO:589, SEQ ID NO:590, SEQ ID NO:591, SEQ ID NO:592, SEQ ID NO:594, SEQ ID NO:604, SEQ ID NO:606, SEQ ID NO:608, SEQ ID NO:610, SEQ ID NO:612, SEQ ID NO:614, SEQ ID NO:616, SEQ ID NO:618, SEQ ID NO:620 or SEQ ID NO:622, and subsequences thereof and variants thereof. In one aspect, the polypeptide has an amylase activity, e.g., an alpha amylase activity or a glucoamylase activity.

The identity can be over the full length of the polypeptide, or, the identity can be over a region of at least about 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 90, 100, 125, 150, 175, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700 or more residues. Polypeptides of the invention can also be shorter than the full length of exemplary polypeptides. In alternative aspects, the invention provides polypeptides (peptides, fragments) ranging in size between about 5 and the full length of a polypeptide, e.g., an enzyme, such as an amylase; exemplary sizes being of about 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 100, 125, 150, 175, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, or more residues, e.g., contiguous residues of an exemplary amylase of the invention. Peptides of the invention can be useful as, e.g., labeling probes, antigens, toleragens, motifs, amylase active sites.

For example, the following table summarizes characteristics (e.g., activity, initial source, signal sequence location and exemplary signal sequence) of exemplary polypeptides of the invention. For example, the polypeptide having a sequence as set forth in SEQ ID NO:437, encoded by SEQ ID NO:436, was artificially generated; the polypeptide having a sequence as set forth in SEQ ID NO:439, encoded by SEQ ID NO:438, has amylase activity under alkaline conditions and was initially derived (isolated) from an unknown source; the polypeptide having a sequence as set forth in SEQ ID NO:441, encoded by SEQ ID NO:440, has amylase activity under alkaline conditions and was initially derived (isolated) from an unknown source, and has a signal sequence consisting of amino acid residues 1 to 32 of SEQ ID NO:441 ("AA 1-32"); see also discussion below regarding signal sequences of the invention, etc.:

SEQ ID NO:P	NOTES	Source	Signal location	Signal Sequence
436, 437	Reassembled amylase	Artificial		
438, 439	ALKALINE AMYLASE	Unknown		

440, 441	ALKALINE AMYLASE	Unknown	AA1-32	MNQIVNFKSHFYRKIALLFSITFI WAAGSLSA
442, 443	ALKALINE AMYLASE	Unknown	AA1-27	MNRYLRLAALTLALAPLAYPWG NLVRA
444, 445	ALKALINE AMYLASE	Unknown	AA1-24	MTPFGQPMMPGARMAAANMA PVRA
446, 447	ALKALINE AMYLASE	Unknown		
448, 449	ALKALINE AMYLASE	Unknown	AA1-23	MRLIMKKMIILITLAWVFTGCES
450, 451	ALKALINE AMYLASE	Unknown	AA1-49	MNDSINLYNFFPYNRPMSINKTN TMKQMINWLGSLALLMLLLSCG EATE
452, 453	ALKALINE AMYLASE	Unknown	AA1-34	MMQLNPWFSTTLKAAGLATALA AVSACQPASESA
454, 455	ALKALINE AMYLASE	Unknown	AA1-37	MDLLEYKNTIQRRQTMTDRKLL FIVATVILAVLVSFS
456, 457	ALKALINE AMYLASE	Unknown	AA1-26	MMQLNPWFSASLKAAGLATALA AVSA
458, 459	ALKALINE AMYLASE	Unknown	AA1-29	MFKVSLRSKDMKKLSLIVTILVLA LTLSA
460, 461	fungal	Cochliobolus heterostrophus ATCC 48331		
462-466	fungal	Fungal	AA1-22	MSRSSTILFVLAAANLASLVDA
467-474	fungal	Cochliobolus heterostrophus ATCC 48331	NOTE: AA1-122 may be removed and the remaining DNA/protein sequences still encode for an amylase	
475-479	fungal	Fungal		
480-485	fungal	Fungal	AA1-19	MKFSLLATIVASISPLARA
486-493	fungal	Fungal	AA1-54	MRRKSTDKYKKVSIRAHLAACE QLAISKMLFSRTATILSLLCVQAT AISPRGSA
494-499	fungal	Fungal	AA1-22	MGFSKMLLGALIGIASLNGVQS
500-510	fungal	Fungal		
511-516	fungal	Fungal	AA1-21	MKYSIIPFVPLFAGLSRAASS
517, 518	fungal	Fungal	AA1-26	MNMNIFLLIISLAFFSTVNCYTMS NA
519-523	fungal	Fungal		
524-528	fungal	Fungal		

529-533	fungal	Cochliobolus heterostrophus ATCC 48331		
534-540	fungal	Cochliobolus heterostrophus ATCC 48331	AA1-20	MLLLNIFTTLFFYITCIVSA
541-545	fungal	Fungal		
546-553	fungal	Fungal	AA1-23	MASSLLSSLSSISTFNSTQILQA
554-559	fungal	Cochliobolus heterostrophus ATCC 48331	AA1-19	MTTALSSGQVAPTPHTAAA
560-566	fungal	Fungal	AA1-33	MLTTSERKTSTAFVTWSMLWVV LLTSFVKDVHA
567, 568	ALKALINE AMYLASE	Unknown		
569, 570		Thermococcus alcaliphilus AEDII12RA		
571, 572		Unknown	AA1-28	MQSNGNVKGRSAVLALALLLLT AVAATA
573, 574		Bacteria	AA1-27	MKKTFKLILVLMLSLTLVFGLTAP IQA
575, 576		Unknown		
577, 578		Unknown	AA1-34	MKPFLKKSIITLLASTCLFTAWLI PSIAVPTVSA
579, 580		Unknown	AA1-29	MFKRRALGFLLAFLLVFTAVFGS MPMEFA
581, 582		Unknown	AA1-27	MKKFYKLTTALALSLSLALSLLG PAHA
583, 584		Unknown		
585, 586		Bacteria	AA1-28	MSLFKKSFPWILSLLLLFLFIAPF SIQT
587-594	GLUCOAMYLASE	Thermomyces lanuginosus ATCC 200065	AA1-23	MLFQPTLCAALGLAALIVQGGEA
603, 604		Unknown	AA1-31	MQNTAKNSIWQRVRHSAIALSA LSLSFGLQA
605, 606		Unknown	AA1-34	MVNHLKKWIAGMALTLALLTGT VVPGLPVQVASA
607, 608		Unknown		

609, 610 611, 612	Unknown Unknown	AA1-31	MQNTAKNSIWQRVRHSAIALSA LSLSFGLQA
613, 614	Unknown	AA1-31	MQNTAKNSIWQRVRHSAIALSA LSLSFGLQA
615, 616	Unknown	AA1-34	MSERGVRRAVRTALVGLAAAAT AAVTLGAPTAQA
617, 618	Unknown	AA1-27	MNRYLRLAALTLALAPLAYPWG NLARA
619, 620	Bacteria	AA1-29	MARKSVAAALALVAGAAAVAVT GNTAAQA
621, 622	Unknown	AA1-31	MQNTAKNSIWQRVRHSAIALSA LSLSFGLQA

Polypeptides and peptides of the invention can be isolated from natural sources, be synthetic, or be recombinantly generated polypeptides. Peptides and proteins can be recombinantly expressed *in vitro* or *in vivo*. The peptides and polypeptides of the invention can be made and isolated using any method known in the art. Polypeptide and peptides of the invention can also be synthesized, whole or in part, using chemical methods well known in the art. See e.g., Caruthers (1980) Nucleic Acids Res. Symp. Ser. 215-223; Horn (1980) Nucleic Acids Res. Symp. Ser. 225-232; Banga, A.K., Therapeutic Peptides and Proteins, Formulation, Processing and Delivery Systems (1995) Technomic Publishing Co., Lancaster, PA. For example, peptide synthesis can be performed using various solid-phase techniques (see e.g., Roberge (1995) Science 269:202; Merrifield (1997) Methods Enzymol. 289:3-13) and automated synthesis may be achieved, e.g., using the ABI 431A Peptide Synthesizer (Perkin Elmer) in accordance with the instructions provided by the manufacturer.

The peptides and polypeptides of the invention can also be glycosylated. The glycosylation can be added post-translationally either chemically or by cellular biosynthetic mechanisms, wherein the later incorporates the use of known glycosylation motifs, which can be native to the sequence or can be added as a peptide or added in the nucleic acid coding sequence. The glycosylation can be O-linked or N-linked. Glycosylation can be added to any polypeptide of the invention to generate an enzyme that is more thermotolerant or thermostable than the "parent" enzyme (to which the

glycosylation was added). The glycosylation can be added by either chemical or by cellular biosynthetic mechanisms.

The invention provides amylases having a broad range of specific activity over a broad range of temperatures, e.g., at about 37°C in the range from about 10 to 10,000, or, 100 to about 1000 units per milligram of protein. Amylases of the invention can also have activity at temperatures as high as 120°C. In alternative aspects, the amylase used in these methods is active at these temperatures, e.g., active at temperatures in a range of between about 80°C to about 115°C, between about 100°C to about 110°C, and from about 105°C to about 108°C. However, amylases of the invention can also have activity at low temperatures, e.g., as low as 4°C to 5°C.

The Tm of an enzyme of the invention can be shifted (for example, can be shifted between about 10°C to 90°C) by heat activation. For example, the Tm of SEQ ID NO:336/337 can be shifted about 17°C to 87°C by heat activation: for example, 80°C preincubation for 5 minutes.

The peptides and polypeptides of the invention, as defined above, include all "mimetic" and "peptidomimetic" forms. The terms "mimetic" and "peptidomimetic" refer to a synthetic chemical compound which has substantially the same structural and/or functional characteristics of the polypeptides of the invention. The mimetic can be either entirely composed of synthetic, non-natural analogues of amino acids, or, is a chimeric molecule of partly natural peptide amino acids and partly non-natural analogs of amino acids. The mimetic can also incorporate any amount of natural amino acid conservative substitutions as long as such substitutions also do not substantially alter the mimetic's structure and/or activity. As with polypeptides of the invention which are conservative variants, routine experimentation will determine whether a mimetic is within the scope of the invention, i.e., that its structure and/or function is not substantially altered. Thus, in one aspect, a mimetic composition is within the scope of the invention if it has an amylase activity.

Polypeptide mimetic compositions of the invention can contain any combination of non-natural structural components. In alternative aspect, mimetic compositions of the invention include one or all of the following three structural groups: a) residue linkage groups other than the natural amide bond ("peptide bond") linkages; b) non-natural residues in place of naturally occurring amino acid residues; or c) residues which induce secondary structural mimicry, i.e., to induce or stabilize a secondary

structure, e.g., a beta turn, gamma turn, beta sheet, alpha helix conformation, and the like. For example, a polypeptide of the invention can be characterized as a mimetic when all or some of its residues are joined by chemical means other than natural peptide bonds. Individual peptidomimetic residues can be joined by peptide bonds, other chemical bonds or coupling means, such as, e.g., glutaraldehyde, N-hydroxysuccinimide esters, bifunctional maleimides, N,N'-dicyclohexylcarbodiimide (DCC) or N,N'-diisopropylcarbodiimide (DIC). Linking groups that can be an alternative to the traditional amide bond ("peptide bond") linkages include, e.g., ketomethylene (e.g., -C(=O)-CH_2- for -C(=O)-NH-), aminomethylene (CH_2-NH), ethylene, olefin (CH=CH), ether (CH_2-O), thioether (CH_2-S), tetrazole (CN_4-), thiazole, retroamide, thioamide, or ester (see, e.g., Spatola (1983) in Chemistry and Biochemistry of Amino Acids, Peptides and Proteins, Vol. 7, pp 267-357, "Peptide Backbone Modifications," Marcell Dekker, NY).

A polypeptide of the invention can also be characterized as a mimetic by containing all or some non-natural residues in place of naturally occurring amino acid residues. Non-natural residues are well described in the scientific and patent literature; a few exemplary non-natural compositions useful as mimetics of natural amino acid residues and guidelines are described below. Mimetics of aromatic amino acids can be generated by replacing by, e.g., D- or L- naphylalanine; D- or L- phenylglycine; D- or L-2 thieneylalanine; D- or L-1, -2, 3-, or 4- pyreneylalanine; D- or L-3 thieneylalanine; D- or L-(2-pyridinyl)-alanine; D- or L-(3-pyridinyl)-alanine; D- or L-(2-pyrazinyl)-alanine; D- or L-(4-isopropyl)-phenylglycine; D-(trifluoromethyl)-phenylglycine; D-(trifluoromethyl)-phenylalanine; D-p-fluoro-phenylalanine; D- or L-p-biphenylphenylalanine; D- or L-p-methoxy-biphenylphenylalanine; D- or L-2-indole(alkyl)alanines; and, D- or L-alkylainines, where alkyl can be substituted or unsubstituted methyl, ethyl, propyl, hexyl, butyl, pentyl, isopropyl, iso-butyl, sec-isotyl, iso-pentyl, or a non-acidic amino acids. Aromatic rings of a non-natural amino acid include, e.g., thiazolyl, thiophenyl, pyrazolyl, benzimidazolyl, naphthyl, furanyl, pyrrolyl, and pyridyl aromatic rings.

Mimetics of acidic amino acids can be generated by substitution by, e.g., non-carboxylate amino acids while maintaining a negative charge; (phosphono)alanine; sulfated threonine. Carboxyl side groups (e.g., aspartyl or glutamyl) can also be selectively modified by reaction with carbodiimides (R'-N-C-N-R') such as, e.g., 1-

cyclohexyl-3(2-morpholinyl-(4-ethyl) carbodiimide or 1-ethyl-3(4-azonia- 4,4-dimetholpentyl) carbodiimide. Aspartyl or glutamyl can also be converted to asparaginyl and glutaminyl residues by reaction with ammonium ions. Mimetics of basic amino acids can be generated by substitution with, e.g., (in addition to lysine and arginine) the amino acids ornithine, citrulline, or (guanidino)-acetic acid, or (guanidino)alkyl-acetic acid, where alkyl is defined above. Nitrile derivative (e.g., containing the CN-moiety in place of COOH) can be substituted for asparagine or glutamine. Asparaginyl and glutaminyl residues can be deaminated to the corresponding aspartyl or glutamyl residues. Arginine residue mimetics can be generated by reacting arginyl with, e.g., one or more conventional reagents, including, e.g., phenylglyoxal, 2,3-butanedione, 1,2-cyclohexanedione, or ninhydrin, preferably under alkaline conditions. Tyrosine residue mimetics can be generated by reacting tyrosyl with, e.g., aromatic diazonium compounds or tetranitromethane. N-acetylimidizol and tetranitromethane can be used to form O-acetyl tyrosyl species and 3-nitro derivatives, respectively. Cysteine residue mimetics can be generated by reacting cysteinyl residues with, e.g., alpha-haloacetates such as 2-chloroacetic acid or chloroacetamide and corresponding amines; to give carboxymethyl or carboxyamidomethyl derivatives. Cysteine residue mimetics can also be generated by reacting cysteinyl residues with, e.g., bromo-trifluoroacetone, alpha-bromo-beta-(5-imidozoyl) propionic acid; chloroacetyl phosphate, N-alkylmaleimides, 3-nitro-2-pyridyl disulfide; methyl 2-pyridyl disulfide; p-chloromercuribenzoate; 2-chloromercuri-4 nitrophenol; or, chloro-7-nitrobenzo-oxa-1,3-diazole. Lysine mimetics can be generated (and amino terminal residues can be altered) by reacting lysinyl with, e.g., succinic or other carboxylic acid anhydrides. Lysine and other alpha-amino-containing residue mimetics can also be generated by reaction with imidoesters, such as methyl picolinimidate, pyridoxal phosphate, pyridoxal, chloroborohydride, trinitro-benzenesulfonic acid, O-methylisourea, 2,4, pentanedione, and transamidase-catalyzed reactions with glyoxylate. Mimetics of methionine can be generated by reaction with, e.g., methionine sulfoxide. Mimetics of proline include, e.g., pipecolic acid, thiazolidine carboxylic acid, 3- or 4- hydroxy proline, dehydroproline, 3- or 4-methylproline, or 3,3,-dimethylproline. Histidine residue mimetics can be generated by reacting histidyl with, e.g., diethylprocarbonate or para-bromophenacyl bromide. Other mimetics include, e.g., those generated by hydroxylation of proline and lysine; phosphorylation of the hydroxyl groups of seryl or threonyl residues; methylation of the alpha-amino groups of lysine,

arginine and histidine; acetylation of the N-terminal amine; methylation of main chain amide residues or substitution with N-methyl amino acids; or amidation of C-terminal carboxyl groups.

A residue, e.g., an amino acid, of a polypeptide of the invention can also be replaced by an amino acid (or peptidomimetic residue) of the opposite chirality. Thus, any amino acid naturally occurring in the L-configuration (which can also be referred to as the R or S, depending upon the structure of the chemical entity) can be replaced with the amino acid of the same chemical structural type or a peptidomimetic, but of the opposite chirality, referred to as the D- amino acid, but also can be referred to as the R- or S- form.

The invention also provides methods for modifying the polypeptides of the invention by either natural processes, such as post-translational processing (e.g., phosphorylation, acylation, etc), or by chemical modification techniques, and the resulting modified polypeptides. Modifications can occur anywhere in the polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. It will be appreciated that the same type of modification may be present in the same or varying degrees at several sites in a given polypeptide. Also a given polypeptide may have many types of modifications. Modifications include acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of a phosphatidylinositol, cross-linking cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cysteine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristolyation, oxidation, pegylation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, and transfer-RNA mediated addition of amino acids to protein such as arginylation. See, e.g., Creighton, T.E., Proteins – Structure and Molecular Properties 2nd Ed., W.H. Freeman and Company, New York (1993); Posttranslational Covalent Modification of Proteins, B.C. Johnson, Ed., Academic Press, New York, pp. 1-12 (1983).

Solid-phase chemical peptide synthesis methods can also be used to synthesize the polypeptide or fragments of the invention. Such method have been known in the art since the early 1960's (Merrifield, R. B., J. Am. Chem. Soc., 85:2149-2154,

1963) (See also Stewart, J. M. and Young, J. D., Solid Phase Peptide Synthesis, 2nd Ed., Pierce Chemical Co., Rockford, Ill., pp. 11-12)) and have recently been employed in commercially available laboratory peptide design and synthesis kits (Cambridge Research Biochemicals). Such commercially available laboratory kits have generally utilized the teachings of H. M. Geysen et al, Proc. Natl. Acad. Sci., USA, 81:3998 (1984) and provide for synthesizing peptides upon the tips of a multitude of "rods" or "pins" all of which are connected to a single plate. When such a system is utilized, a plate of rods or pins is inverted and inserted into a second plate of corresponding wells or reservoirs, which contain solutions for attaching or anchoring an appropriate amino acid to the pin's or rod's tips. By repeating such a process step, i.e., inverting and inserting the rod's and pin's tips into appropriate solutions, amino acids are built into desired peptides. In addition, a number of available FMOC peptide synthesis systems are available. For example, assembly of a polypeptide or fragment can be carried out on a solid support using an Applied Biosystems, Inc. Model 431A™ automated peptide synthesizer. Such equipment provides ready access to the peptides of the invention, either by direct synthesis or by synthesis of a series of fragments that can be coupled using other known techniques.

The invention provides novel amylases (e.g., alpha amylases), including the exemplary enzymes of the invention, nucleic acids encoding them, antibodies that bind them, and methods for making and using them. In one aspect, the polypeptides of the invention have an amylase activity, as described herein, including, e.g., the ability to hydrolyze starches into sugars. In one aspect, the polypeptides of the invention have an alpha amylase activity. In alternative aspects, the amylases of the invention have activities that have been modified from those of the exemplary amylases described herein.

The invention includes amylases of the invention with and without signal sequences (including signal sequences of the invention, see e.g., Table 3, below, or other signal sequences) and the signal sequences themselves (e.g., Table 3, below). The invention also include polypeptides (e.g., fusion proteins) comprising a signal sequence of the invention, see, e.g., Table 3, below. The polypeptide comprising a signal sequence of the invention can be an amylase of the invention or another amylase or another enzyme or other polypeptide.

The invention includes immobilized amylases, anti-amylase antibodies and fragments thereof. The invention provides methods for inhibiting amylase activity, e.g., using dominant negative mutants or anti-amylase antibodies of the invention. The

invention includes heterocomplexes, e.g., fusion proteins, heterodimers, etc., comprising the amylases of the invention.

In one aspect, amylases (e.g., alpha amylases) of the invention hydrolyze internal polysaccharide bonds, e.g., a-1,4- and 1,6-glucosidic bonds in starch to produce smaller molecular weight maltodextrines. In one aspect, this hydrolysis is largely at random. Thus, the invention provides methods for producing smaller molecular weight maltodextrines.

Amylases of the invention can be used in laboratory and industrial settings to hydrolyze starch or any maltodextrine-comprising compound for a variety of purposes. These amylases can be used alone to provide specific hydrolysis or can be combined with other amylases to provide a "cocktail" with a broad spectrum of activity. Exemplary uses include the removal or partial or complete hydrolysis of starch or any maltodextrine-comprising compound from biological, food, animal feed, pharmaceutical or industrial samples.

For example, the amylases of the present invention can be formulated in laundry detergents to aid in the removal of starch-containing stains. In one aspect, the invention provides detergents comprising amylases of the invention, including amylases active under alkaline conditions, and methods of making and using them. These detergent compositions include laundry and dishwashing (e.g., autodishwashing) solutions and application. Amylases of the invention can be used as cleaning agents in any detergent matrices (see industrial applications below). The amylases of the present invention can be used in the initial stages (liquefaction) of starch processing, in wet corn milling, in alcohol production, in the textile industry for starch desizing, in baking applications, in the beverage industry, in oilfields in drilling processes; in inking of recycled paper; and in animal feed.

Amylases of the invention can have an amylase activity under various conditions, e.g., extremes in pH and/or temperature, oxidizing agents, and the like. The invention provides methods leading to alternative amylase preparations with different catalytic efficiencies and stabilities, e.g., towards temperature, oxidizing agents and changing wash conditions. In one aspect, amylase variants can be produced using techniques of site-directed mutagenesis and/or random mutagenesis. In one aspect, directed evolution can be used to produce a great variety of amylase variants with alternative specificities and stability.

The proteins of the invention are also useful as research reagents to identify amylase modulators, e.g., activators or inhibitors of amylase activity. Briefly, test samples (compounds, broths, extracts, and the like) are added to amylase assays to determine their ability to inhibit substrate cleavage. Inhibitors identified in this way can be used in industry and research to reduce or prevent undesired proteolysis. As with amylases, inhibitors can be combined to increase the spectrum of activity.

The invention also provides methods of discovering new amylases using the nucleic acids, polypeptides and antibodies of the invention. In one aspect, lambda phage libraries are screened for expression-based discovery of amylases. In one aspect, the invention uses lambda phage libraries in screening to allow detection of toxic clones; improved access to substrate; reduced need for engineering a host, by-passing the potential for any bias resulting from mass excision of the library; and, faster growth at low clone densities. Screening of lambda phage libraries can be in liquid phase or in solid phase. In one aspect, the invention provides screening in liquid phase. This gives a greater flexibility in assay conditions; additional substrate flexibility; higher sensitivity for weak clones; and ease of automation over solid phase screening.

The invention provides screening methods using the proteins and nucleic acids of the invention and robotic automation to enable the execution of many thousands of biocatalytic reactions and screening assays in a short period of time, e.g., per day, as well as ensuring a high level of accuracy and reproducibility (see discussion of arrays, below). As a result, a library of derivative compounds can be produced in a matter of weeks. For further teachings on modification of molecules, including small molecules, see PCT/US94/09174.

The present invention includes amylase enzymes which are non-naturally occurring carbonyl hydrolase variants (e.g., amylase variants) having a different proteolytic activity, stability, substrate specificity, pH profile and/or performance characteristic as compared to the precursor carbonyl hydrolase from which the amino acid sequence of the variant is derived. Specifically, such amylase variants have an amino acid sequence not found in nature, which is derived by substitution of a plurality of amino acid residues of a precursor amylase with different amino acids. The precursor amylase may be a naturally-occurring amylase or a recombinant amylase. The useful amylase variants encompass the substitution of any of the naturally occurring L-amino acids at the designated amino acid residue positions.

Amylase Signal Sequences

The invention provides signal sequences consisting of or comprising a peptide having a sequence comprising residues 1 to 12, 1 to 13, 1 to 14, 1 to 15, 1 to 16, 1 to 17, 1 to 18, 1 to 19, 1 to 20, 1 to 21, 1 to 22, 1 to 23, 1 to 24, 1 to 25, 1 to 26, 1 to 27, 1 to 28, 1 to 28, 1 to 30 or 1 to 31, 1 to 32, 1 to 33, 1 to 34, 1 to 35, 1 to 36, 1 to 37, 1 to 38, or 1 to 39, or longer, of a polypeptide of the invention. For example, the invention provides amylase (e.g., alpha amylase or glucoamylase) signal sequences and nucleic acids encoding these signal sequences, e.g., exemplary peptides of the invention having sequences as set forth in Table 3, SEQ ID NO:7, SEQ ID NO:8, and SEQ ID NO:213 through 257, and polypeptides comprising (or consisting of) sequences as set forth in Table 3, SEQ ID NO:7, SEQ ID NO:8, and SEQ ID NO:213 through 257. The invention also provides amylase signal sequences and nucleic acids encoding these signal sequences, e.g., peptides comprising or consisting of residues 1 to 27 of SEQ ID NO:323 (encoded by SEQ ID NO:322), peptides comprising or consisting of residues 1 to 22 of SEQ ID NO:333 (encoded by SEQ ID NO:332), peptides comprising or consisting of residues 1 to 20 of SEQ ID NO:335 (encoded by SEQ ID NO:334), peptides comprising or consisting of residues 1 to 35 of SEQ ID NO:337 (encoded by SEQ ID NO:336), etc., see Table 3 for, in addition to these signal sequences, additional amylase signal sequences and nucleic acids encoding these signal sequences.

The invention also provides amylase signal sequences and nucleic acids encoding these signal sequences comprising or consisting of residues 1 to 32 or 1 to 33 of SEQ ID NO:441; residues 1 to 27 or 1 to 28 of SEQ ID NO:443; residues 1 to 24 or 1 to 25 of SEQ ID NO:445; residues 1 to 23 or 1 to 24 of SEQ ID NO:449; residues 1 to 49 or 1 to 50 of SEQ ID NO:451; residues 1 to 34 or 1 to 35 of SEQ ID NO:453; residues 1 to 37 or 1 to 38 of SEQ ID NO:455; residues 1 to 26 or 1 to 27 of SEQ ID NO:457; residues 1 to 29 or 1 to 30 of SEQ ID NO:459; residues 1 to 22 or 1 to 23 of SEQ ID NO:466; residues 1 to 19 or 1 to 20 of SEQ ID NO:485; residues 1 to 54 or 1 to 55 of SEQ ID NO:493; residues 1 to 22 to 1 to 23 of SEQ ID NO:499; residues 21 or 1 to 22 of SEQ ID NO:516; residues 1 to 26 or 1 to 27 of SEQ ID NO:518; residues 1 to 20 or 1 to 21 of SEQ ID NO:540; residues 1 to 23 or 1 to 24 of SEQ ID NO:553; residues 1 to 19 or 1 to 20 of SEQ ID NO:559; residues 1 to 33 or 1 to 34 of SEQ ID NO:566.

For example, regarding Table 3, the invention provides peptides comprising or consisting of amino acid residues 1 to 23 (SEQ ID NO:213) of SEQ ID NO:87, etc.

Table 3

SEQ ID NO.	Signal Sequence
SEQ ID NO: 87	AA1-23 (SEQ ID NO:213)
SEQ ID NO: 91	AA1-23 (SEQ ID NO: 214)
SEQ ID NO: 93	AA1-33 (SEQ ID NO: 215)
SEQ ID NO: 97	AA1-31 (SEQ ID NO: 216)
SEQ ID NO: 99	AA1-30 (SEQ ID NO: 217)
SEQ ID NO: 103	AA1-22 (SEQ ID NO: 218)
SEQ ID NO: 105	AA1-33 (SEQ ID NO: 219)
SEQ ID NO: 109	AA1-25 (SEQ ID NO: 220)
SEQ ID NO: 111	AA1-35 (SEQ ID NO: 221)
SEQ ID NO: 113	AA1-28 (SEQ ID NO: 222)
SEQ ID NO: 117	AA1-21 (SEQ ID NO: 223)
SEQ ID NO: 119	AA1-30 (SEQ ID NO: 224)
SEQ ID NO: 123	AA1-35 (SEQ ID NO: 225)
SEQ ID NO: 125	AA1-28 (SEQ ID NO: 226)
SEQ ID NO: 127	AA1-30 (SEQ ID NO: 227)
SEQ ID NO: 131	AA1-30 (SEQ ID NO: 228)
SEQ ID NO: 133	AA1-30 (SEQ ID NO: 229)
SEQ ID NO: 137	AA1-28 (SEQ ID NO: 230)
SEQ ID NO: 139	AA1-23 (SEQ ID NO: 231)
SEQ ID NO: 141	AA1-23 (SEQ ID NO: 232)
SEQ ID NO: 143	AA1-30 (SEQ ID NO: 233)
SEQ ID NO: 145	AA1-27 (SEQ ID NO: 234)
SEQ ID NO: 147	AA1-29 (SEQ ID NO: 235)
SEQ ID NO: 149	AA1-28 (SEQ ID NO: 236)
SEQ ID NO: 69	AA1-27 (SEQ ID NO: 237)
SEQ ID NO: 153	AA1-26 (SEQ ID NO: 238)
SEQ ID NO: 155	AA1-33 (SEQ ID NO: 239)
SEQ ID NO: 157	AA1-25 (SEQ ID NO: 240)
SEQ ID NO: 159	AA1-25 (SEQ ID NO: 241)
SEQ ID NO: 161	AA1-36 (SEQ ID NO: 242)
SEQ ID NO: 167	AA1-36 (SEQ ID NO: 243)
SEQ ID NO: 169	AA1-23 (SEQ ID NO: 244)
SEQ ID NO: 173	AA1-25 (SEQ ID NO: 245)

SEQ ID NO: 175	AA1-22 (SEQ ID NO: 246)
SEQ ID NO: 177	AA1-23 (SEQ ID NO: 247)
SEQ ID NO: 179	AA1-23 (SEQ ID NO: 248)
SEQ ID NO: 185	AA1-25 (SEQ ID NO: 249)
SEQ ID NO: 189	AA1-36 (SEQ ID NO: 250)
SEQ ID NO: 191	AA1-25 (SEQ ID NO: 251)
SEQ ID NO: 193	AA1-25 (SEQ ID NO: 252)
SEQ ID NO: 197	AA1-23 (SEQ ID NO: 253)
SEQ ID NO: 199	AA1-23 (SEQ ID NO: 254)
SEQ ID NO: 201	AA1-30 (SEQ ID NO: 255)
SEQ ID NO: 203	AA1-25 (SEQ ID NO: 256)
SEQ ID NO: 205	AA1-16 (SEQ ID NO: 257)
SEQ ID NO:73	AA1-16 (SEQ ID NO: 7)
SEQ ID NO: 79	AA1-26 (SEQ ID NO: 8)
SEQ ID NO: 322, 323	Residues 1 through 27
SEQ ID NO: 332, 333	Residues 1 through 22
SEQ ID NO:334, 335	Residues 1 through 20
SEQ ID NO:336, 337	Residues 1 through 35
SEQ ID NO:338, 339	Residues 1 through 50
SEQ ID NO:342, 343	Residues 1 through 23
SEQ ID NO:344, 345	Residues 1 through 22
SEQ ID NO:346, 347	Residues 1 through 21
SEQ ID NO:350, 351	Residues 1 through 21
SEQ ID NO:352, 353	Residues 1 through 27
SEQ ID NO:354, 355	Residues 1 through 24
SEQ ID NO:358, 359	Residues 1 through 29
SEQ ID NO:362, 363	Residues 1 through 20
SEQ ID NO:364, 365	Residues 1 through 29
SEQ ID NO:366, 367	Residues 1 through 24
SEQ ID NO:370, 371	Residues 1 through 22
SEQ ID NO:372, 373	Residues 1 through 25
SEQ ID NO:374, 375	Residues 1 through 21
SEQ ID NO:376, 377	Residues 1 through 37
SEQ ID NO:378, 379	Residues 1 through 27
SEQ ID NO:380, 381	Residues 1 through 29
SEQ ID NO:382, 383	Residues 1 through 35
SEQ ID NO:384, 385	Residues 1 through 37
SEQ ID NO:386, 387	Residues 1 through 25
SEQ ID NO:388, 389	Residues 1 through 21

SEQ ID NO:390, 391	Residues 1 through 58
SEQ ID NO:394, 395	Residues 1 through 57
SEQ ID NO:396, 397	Residues 1 through 19
SEQ ID NO:400, 401	Residues 1 through 19
SEQ ID NO:402, 403	Residues 1 through 19
SEQ ID NO:404, 405	Residues 1 through 26
SEQ ID NO:406, 407	Residues 1 through 21
SEQ ID NO:408, 409	Residues 1 through 51
SEQ ID NO:410, 411	Residues 1 through 21
SEQ ID NO:416, 417	Residues 1 through 24
SEQ ID NO:418, 419	Residues 1 through 44
SEQ ID NO:420, 421	Residues 1 through 44
SEQ ID NO:422, 423	Residues 1 through 27
SEQ ID NO:424, 425	Residues 1 through 37
SEQ ID NO:428, 429	Residues 1 through 30
SEQ ID NO:430, 431	Residues 1 through 33
SEQ ID NO:432, 433	Residues 1 through 34
SEQ ID NO:434, 435	Residues 1 through 27

The amylase signal sequences of the invention can be isolated peptides, or, sequences joined to another amylase or a non-amylase polypeptide, e.g., as a fusion protein. In one aspect, the invention provides polypeptides comprising amylase signal sequences of the invention. In one aspect, polypeptides comprising amylase signal sequences of the invention comprise sequences heterologous to an amylase of the invention (e.g., a fusion protein comprising an amylase signal sequence of the invention and sequences from another amylase or a non-amylase protein). In one aspect, the invention provides amylases of the invention with heterologous signal sequences, e.g., sequences with a yeast signal sequence. For example, an amylase of the invention comprising a heterologous signal sequence in a vectors, e.g., a pPIC series vector (Invitrogen, Carlsbad, CA).

In one aspect, the signal sequences of the invention are identified following identification of novel amylase polypeptides. The pathways by which proteins are sorted and transported to their proper cellular location are often referred to as protein targeting pathways. One of the most important elements in all of these targeting systems is a short amino acid sequence at the amino terminus of a newly synthesized polypeptide called the signal sequence. This signal sequence directs a protein to its appropriate

location in the cell and is removed during transport or when the protein reaches its final destination. Most lysosomal, membrane, or secreted proteins have an amino-terminal signal sequence that marks them for translocation into the lumen of the endoplasmic reticulum. More than 100 signal sequences for proteins in this group have been determined. The signal sequences can vary in length from 13 to 36 amino acid residues. Various methods of recognition of signal sequences are known to those of skill in the art. For example, in one aspect, novel amylase signal peptides are identified by a method referred to as SignalP. SignalP uses a combined neural network which recognizes both signal peptides and their cleavage sites. (Nielsen, et al., "Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites." Protein Engineering, vol. 10, no. 1, p. 1-6 (1997).

It should be understood that in some aspects amylases of the invention may not have signal sequences. In one aspect, the invention provides the amylases of the invention lacking all or part of a signal sequence, e.g. the signal sequences of the invention (see Table 3, below). In one aspect, the invention provides a nucleic acid sequence encoding a signal sequence from one amylase operably linked to a nucleic acid sequence of a different amylase or, optionally, a signal sequence from a non-amylase protein may be desired. Table 3 shows exemplary signal sequences of the invention.

Amylase prepro and signal sequences and catalytic domains

In addition to signal sequences (e.g., signal peptides (SPs)), as discussed above, the invention provides prepro domains and catalytic domains (CDs). The SPs, prepro domains and/or CDs of the invention can be isolated or recombinant peptides or can be part of a fusion protein, e.g., as a heterologous domain in a chimeric protein. The invention provides nucleic acids encoding these catalytic domains (CDs) (e.g., "active sites"), prepro domains and signal sequences (SPs, e.g., a peptide having a sequence comprising/ consisting of amino terminal residues of a polypeptide of the invention).

The amylase signal sequences (SPs), catalytic domains (CDs) and/or prepro sequences of the invention can be isolated peptides, or, sequences joined to another amylase or a non- amylase polypeptide, e.g., as a fusion (chimeric) protein. In one aspect, polypeptides comprising amylase signal sequences SPs and/or prepro of the

invention comprise sequences heterologous to amylases of the invention (e.g., a fusion protein comprising an SP and/or prepro of the invention and sequences from another amylase or a non- amylase protein). In one aspect, the invention provides amylases of the invention with heterologous CDs, SPs and/or prepro sequences, e.g., sequences with a yeast signal sequence. An amylase of the invention can comprise a heterologous CD, SP and/or prepro in a vector, e.g., a pPIC series vector (Invitrogen, Carlsbad, CA).

In one aspect, SPs, CDs, and/or prepro sequences of the invention are identified following identification of novel amylase polypeptides. The pathways by which proteins are sorted and transported to their proper cellular location are often referred to as protein targeting pathways. One of the most important elements in all of these targeting systems is a short amino acid sequence at the amino terminus of a newly synthesized polypeptide called the signal sequence. This signal sequence directs a protein to its appropriate location in the cell and is removed during transport or when the protein reaches its final destination. Most lysosomal, membrane, or secreted proteins have an amino-terminal signal sequence that marks them for translocation into the lumen of the endoplasmic reticulum. The signal sequences can vary in length from 13 to 45 or more amino acid residues. Various methods of recognition of signal sequences are known to those of skill in the art. For example, in one aspect, novel hydrolase signal peptides are identified by a method referred to as SignalP. SignalP uses a combined neural network which recognizes both signal peptides and their cleavage sites. (Nielsen, et al., "Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites." Protein Engineering, vol. 10, no. 1, p. 1-6 (1997).

In some aspects, an amylase of the invention may not have SPs and/or prepro sequences, and/or catalytic domains (CDs). In one aspect, the invention provides amylases lacking all or part of an SP, a CD and/or a prepro domain. In one aspect, the invention provides a nucleic acid sequence encoding a signal sequence (SP), a CD and/or prepro from one amylase operably linked to a nucleic acid sequence of a different amylase or, optionally, a signal sequence (SPs), a CD and/or prepro domain from a non-amylase protein may be desired.

The invention also provides isolated or recombinant polypeptides comprising signal sequences (SPs), prepro domain and/or catalytic domains (CDs) of the invention and heterologous sequences. The heterologous sequences are sequences not naturally associated (e.g., to an amylase) with an SP, prepro domain and/or CD. The

sequence to which the SP, prepro domain and/or CD are not naturally associated can be on the SP's, prepro domain and/or CD's amino terminal end, carboxy terminal end, and/or on both ends of the SP and/or CD. In one aspect, the invention provides an isolated or recombinant polypeptide comprising (or consisting of) a polypeptide comprising a signal sequence (SP), prepro domain and/or catalytic domain (CD) of the invention with the proviso that it is not associated with any sequence to which it is naturally associated (e.g., amylase sequence). Similarly in one aspect, the invention provides isolated or recombinant nucleic acids encoding these polypeptides. Thus, in one aspect, the isolated or recombinant nucleic acid of the invention comprises coding sequence for a signal sequence (SP), prepro domain and/or catalytic domain (CD) of the invention and a heterologous sequence (i.e., a sequence not naturally associated with the a signal sequence (SP), prepro domain and/or catalytic domain (CD) of the invention). The heterologous sequence can be on the 3' terminal end, 5' terminal end, and/or on both ends of the SP, prepro domain and/or CD coding sequence.

The polypeptides of the invention include amylases in an active or inactive form. For example, the polypeptides of the invention include proproteins before "maturation" or processing of prepro sequences, e.g., by a proprotein-processing enzyme, such as a proprotein convertase to generate an "active" mature protein. The polypeptides of the invention include amylases inactive for other reasons, e.g., before "activation" by a post-translational processing event, e.g., an endo- or exo-peptidase or proteinase action, a phosphorylation event, an amidation, a glycosylation or a sulfation, a dimerization event, and the like. Methods for identifying "prepro" domain sequences, CDs, and signal sequences are well known in the art, see, e.g., Van de Ven (1993) Crit. Rev. Oncog. 4(2):115-136. For example, to identify a prepro sequence, the protein is purified from the extracellular space and the N-terminal protein sequence is determined and compared to the unprocessed form.

The polypeptides of the invention include all active forms, including active subsequences, e.g., catalytic domains (CDs) or active sites, of an enzyme of the invention. In one aspect, the invention provides catalytic domains or active sites as set forth below. In one aspect, the invention provides a peptide or polypeptide comprising or consisting of an active site domain as predicted through use of a database such as Pfam (which is a large collection of multiple sequence alignments and hidden Markov models covering many common protein families, The Pfam protein families database, A. Bateman, E.

Birney, L. Cerruti, R. Durbin, L. Etwiller, S.R. Eddy, S. Griffiths-Jones, K.L. Howe, M. Marshall, and E.L.L. Sonnhammer, Nucleic Acids Research, 30(1):276-280, 2002) or equivalent.

Hybrid amylases and peptide libraries

In one aspect, the invention provides hybrid amylases and fusion proteins, including peptide libraries, comprising sequences of the invention. The peptide libraries of the invention can be used to isolate peptide modulators (e.g., activators or inhibitors) of targets, such as amylase substrates, receptors, enzymes. The peptide libraries of the invention can be used to identify formal binding partners of targets, such as ligands, e.g., cytokines, hormones and the like.

In one aspect, the fusion proteins of the invention (e.g., the peptide moiety) are conformationally stabilized (relative to linear peptides) to allow a higher binding affinity for targets. The invention provides fusions of amylases of the invention and other peptides, including known and random peptides. They can be fused in such a manner that the structure of the amylases is not significantly perturbed and the peptide is metabolically or structurally conformationally stabilized. This allows the creation of a peptide library that is easily monitored both for its presence within cells and its quantity.

Amino acid sequence variants of the invention can be characterized by a predetermined nature of the variation, a feature that sets them apart from a naturally occurring form, e.g., an allelic or interspecies variation of an amylase sequence. In one aspect, the variants of the invention exhibit the same qualitative biological activity as the naturally occurring analogue. Alternatively, the variants can be selected for having modified characteristics. In one aspect, while the site or region for introducing an amino acid sequence variation is predetermined, the mutation per se need not be predetermined. For example, in order to optimize the performance of a mutation at a given site, random mutagenesis may be conducted at the target codon or region and the expressed amylase variants screened for the optimal combination of desired activity. Techniques for making substitution mutations at predetermined sites in DNA having a known sequence are well known, as discussed herein for example, M13 primer mutagenesis and PCR mutagenesis. Screening of the mutants can be done using assays of proteolytic activities. In alternative aspects, amino acid substitutions can be single residues; insertions can be on the order of from about 1 to 20 amino acids, although considerably larger insertions can be done.

Deletions can range from about 1 to about 20, 30, 40, 50, 60, 70 residues or more. To obtain a final derivative with the optimal properties, substitutions, deletions, insertions or any combination thereof may be used. Generally, these changes are done on a few amino acids to minimize the alteration of the molecule. However, larger changes may be tolerated in certain circumstances.

The invention provides amylases where the structure of the polypeptide backbone, the secondary or the tertiary structure, e.g., an alpha-helical or beta-sheet structure, has been modified. In one aspect, the charge or hydrophobicity has been modified. In one aspect, the bulk of a side chain has been modified. Substantial changes in function or immunological identity are made by selecting substitutions that are less conservative. For example, substitutions can be made which more significantly affect: the structure of the polypeptide backbone in the area of the alteration, for example a alpha-helical or a beta-sheet structure; a charge or a hydrophobic site of the molecule, which can be at an active site; or a side chain. The invention provides substitutions in polypeptide of the invention where (a) a hydrophilic residues, e.g. seryl or threonyl, is substituted for (or by) a hydrophobic residue, e.g. leucyl, isoleucyl, phenylalanyl, valyl or alanyl; (b) a cysteine or proline is substituted for (or by) any other residue; (c) a residue having an electropositive side chain, e.g. lysyl, arginyl, or histidyl, is substituted for (or by) an electronegative residue, e.g. glutamyl or aspartyl; or (d) a residue having a bulky side chain, e.g. phenylalanine, is substituted for (or by) one not having a side chain, e.g. glycine. The variants can exhibit the same qualitative biological activity (i.e. amylase activity) although variants can be selected to modify the characteristics of the amylases as needed.

In one aspect, amylases of the invention comprise epitopes or purification tags, signal sequences or other fusion sequences, etc. In one aspect, the amylases of the invention can be fused to a random peptide to form a fusion polypeptide. By "fused" or "operably linked" herein is meant that the random peptide and the amylase are linked together, in such a manner as to minimize the disruption to the stability of the amylase structure, e.g., it retains amylase activity. The fusion polypeptide (or fusion polynucleotide encoding the fusion polypeptide) can comprise further components as well, including multiple peptides at multiple loops.

In one aspect, the peptides and nucleic acids encoding them are randomized, either fully randomized or they are biased in their randomization, e.g. in

nucleotide/residue frequency generally or per position. "Randomized" means that each nucleic acid and peptide consists of essentially random nucleotides and amino acids, respectively. In one aspect, the nucleic acids which give rise to the peptides can be chemically synthesized, and thus may incorporate any nucleotide at any position. Thus, when the nucleic acids are expressed to form peptides, any amino acid residue may be incorporated at any position. The synthetic process can be designed to generate randomized nucleic acids, to allow the formation of all or most of the possible combinations over the length of the nucleic acid, thus forming a library of randomized nucleic acids. The library can provide a sufficiently structurally diverse population of randomized expression products to affect a probabilistically sufficient range of cellular responses to provide one or more cells exhibiting a desired response. Thus, the invention provides an interaction library large enough so that at least one of its members will have a structure that gives it affinity for some molecule, protein, or other factor.

Screening Methodologies and "On-line" Monitoring Devices

In practicing the methods of the invention, a variety of apparatus and methodologies can be used to in conjunction with the polypeptides and nucleic acids of the invention, e.g., to screen polypeptides for amylase activity, to screen compounds as potential modulators, e.g., activators or inhibitors, of an amylase activity, for antibodies that bind to a polypeptide of the invention, for nucleic acids that hybridize to a nucleic acid of the invention, to screen for cells expressing a polypeptide of the invention and the like.

Capillary Arrays

Capillary arrays, such as the GIGAMATRIX™, Diversa Corporation, San Diego, CA, can be used to in the methods of the invention. Nucleic acids or polypeptides of the invention can be immobilized to or applied to an array, including capillary arrays. Arrays can be used to screen for or monitor libraries of compositions (e.g., small molecules, antibodies, nucleic acids, etc.) for their ability to bind to or modulate the activity of a nucleic acid or a polypeptide of the invention. Capillary arrays provide another system for holding and screening samples. For example, a sample screening apparatus can include a plurality of capillaries formed into an array of adjacent capillaries, wherein each capillary comprises at least one wall defining a lumen for retaining a sample. The apparatus can further include interstitial material disposed

between adjacent capillaries in the array, and one or more reference indicia formed within of the interstitial material. A capillary for screening a sample, wherein the capillary is adapted for being bound in an array of capillaries, can include a first wall defining a lumen for retaining the sample, and a second wall formed of a filtering material, for filtering excitation energy provided to the lumen to excite the sample.

A polypeptide or nucleic acid, e.g., a ligand, can be introduced into a first component into at least a portion of a capillary of a capillary array. Each capillary of the capillary array can comprise at least one wall defining a lumen for retaining the first component. An air bubble can be introduced into the capillary behind the first component. A second component can be introduced into the capillary, wherein the second component is separated from the first component by the air bubble. A sample of interest can be introduced as a first liquid labeled with a detectable particle into a capillary of a capillary array, wherein each capillary of the capillary array comprises at least one wall defining a lumen for retaining the first liquid and the detectable particle, and wherein the at least one wall is coated with a binding material for binding the detectable particle to the at least one wall. The method can further include removing the first liquid from the capillary tube, wherein the bound detectable particle is maintained within the capillary, and introducing a second liquid into the capillary tube.

The capillary array can include a plurality of individual capillaries comprising at least one outer wall defining a lumen. The outer wall of the capillary can be one or more walls fused together. Similarly, the wall can define a lumen that is cylindrical, square, hexagonal or any other geometric shape so long as the walls form a lumen for retention of a liquid or sample. The capillaries of the capillary array can be held together in close proximity to form a planar structure. The capillaries can be bound together, by being fused (e.g., where the capillaries are made of glass), glued, bonded, or clamped side-by-side. The capillary array can be formed of any number of individual capillaries, for example, a range from 100 to 4,000,000 capillaries. A capillary array can form a micro titer plate having about 100,000 or more individual capillaries bound together.

Arrays, or "Biochips"

Nucleic acids or polypeptides of the invention can be immobilized to or applied to an array. Arrays can be used to screen for or monitor libraries of compositions (e.g., small molecules, antibodies, nucleic acids, etc.) for their ability to bind to or modulate the activity of a nucleic acid or a polypeptide of the invention. For example, in

one aspect of the invention, a monitored parameter is transcript expression of an amylase gene. One or more, or, all the transcripts of a cell can be measured by hybridization of a sample comprising transcripts of the cell, or, nucleic acids representative of or complementary to transcripts of a cell, by hybridization to immobilized nucleic acids on an array, or "biochip." By using an "array" of nucleic acids on a microchip, some or all of the transcripts of a cell can be simultaneously quantified. Alternatively, arrays comprising genomic nucleic acid can also be used to determine the genotype of a newly engineered strain made by the methods of the invention. Polypeptide arrays" can also be used to simultaneously quantify a plurality of proteins. The present invention can be practiced with any known "array," also referred to as a "microarray" or "nucleic acid array" or "polypeptide array" or "antibody array" or "biochip," or variation thereof. Arrays are generically a plurality of "spots" or "target elements," each target element comprising a defined amount of one or more biological molecules, e.g., oligonucleotides, immobilized onto a defined area of a substrate surface for specific binding to a sample molecule, e.g., mRNA transcripts.

In practicing the methods of the invention, any known array and/or method of making and using arrays can be incorporated in whole or in part, or variations thereof, as described, for example, in U.S. Patent Nos. 6,277,628; 6,277,489; 6,261,776; 6,258,606; 6,054,270; 6,048,695; 6,045,996; 6,022,963; 6,013,440; 5,965,452; 5,959,098; 5,856,174; 5,830,645; 5,770,456; 5,632,957; 5,556,752; 5,143,854; 5,807,522; 5,800,992; 5,744,305; 5,700,637; 5,556,752; 5,434,049; see also, e.g., WO 99/51773; WO 99/09217; WO 97/46313; WO 96/17958; see also, e.g., Johnston (1998) Curr. Biol. 8:R171-R174; Schummer (1997) Biotechniques 23:1087-1092; Kern (1997) Biotechniques 23:120-124; Solinas-Toldo (1997) Genes, Chromosomes & Cancer 20:399-407; Bowtell (1999) Nature Genetics Supp. 21:25-32. See also published U.S. patent applications Nos. 20010018642; 20010019827; 20010016322; 20010014449; 20010014448; 20010012537; 20010008765.

Antibodies and Antibody-based screening methods

The invention provides isolated or recombinant antibodies that specifically bind to an amylase of the invention. These antibodies can be used to isolate, identify or quantify the amylases of the invention or related polypeptides. These antibodies can be used to isolate other polypeptides within the scope the invention or other related

amylases. The antibodies can be designed to bind to an active site of an amylase. Thus, the invention provides methods of inhibiting amylases using the antibodies of the invention.

The antibodies can be used in immunoprecipitation, staining, immunoaffinity columns, and the like. If desired, nucleic acid sequences encoding for specific antigens can be generated by immunization followed by isolation of polypeptide or nucleic acid, amplification or cloning and immobilization of polypeptide onto an array of the invention. Alternatively, the methods of the invention can be used to modify the structure of an antibody produced by a cell to be modified, e.g., an antibody's affinity can be increased or decreased. Furthermore, the ability to make or modify antibodies can be a phenotype engineered into a cell by the methods of the invention.

Methods of immunization, producing and isolating antibodies (polyclonal and monoclonal) are known to those of skill in the art and described in the scientific and patent literature, see, e.g., Coligan, CURRENT PROTOCOLS IN IMMUNOLOGY, Wiley/Greene, NY (1991); Stites (eds.) BASIC AND CLINICAL IMMUNOLOGY (7th ed.) Lange Medical Publications, Los Altos, CA ("Stites"); Goding, MONOCLONAL ANTIBODIES: PRINCIPLES AND PRACTICE (2d ed.) Academic Press, New York, NY (1986); Kohler (1975) Nature 256:495; Harlow (1988) ANTIBODIES, A LABORATORY MANUAL, Cold Spring Harbor Publications, New York. Antibodies also can be generated in vitro, e.g., using recombinant antibody binding site expressing phage display libraries, in addition to the traditional in vivo methods using animals. See, e.g., Hoogenboom (1997) Trends Biotechnol. 15:62-70; Katz (1997) Annu. Rev. Biophys. Biomol. Struct. 26:27-45.

Polypeptides or peptides can be used to generate antibodies which bind specifically to the polypeptides, e.g., the amylases, of the invention. The resulting antibodies may be used in immunoaffinity chromatography procedures to isolate or purify the polypeptide or to determine whether the polypeptide is present in a biological sample. In such procedures, a protein preparation, such as an extract, or a biological sample is contacted with an antibody capable of specifically binding to one of the polypeptides of the invention.

In immunoaffinity procedures, the antibody is attached to a solid support, such as a bead or other column matrix. The protein preparation is placed in contact with the antibody under conditions in which the antibody specifically binds to one of the

polypeptides of the invention. After a wash to remove non-specifically bound proteins, the specifically bound polypeptides are eluted.

The ability of proteins in a biological sample to bind to the antibody may be determined using any of a variety of procedures familiar to those skilled in the art. For example, binding may be determined by labeling the antibody with a detectable label such as a fluorescent agent, an enzymatic label, or a radioisotope. Alternatively, binding of the antibody to the sample may be detected using a secondary antibody having such a detectable label thereon. Particular assays include ELISA assays, sandwich assays, radioimmunoassays, and Western Blots.

Polyclonal antibodies generated against the polypeptides of the invention can be obtained by direct injection of the polypeptides into an animal or by administering the polypeptides to a non-human animal. The antibody so obtained will then bind the polypeptide itself. In this manner, even a sequence encoding only a fragment of the polypeptide can be used to generate antibodies which may bind to the whole native polypeptide. Such antibodies can then be used to isolate the polypeptide from cells expressing that polypeptide.

For preparation of monoclonal antibodies, any technique which provides antibodies produced by continuous cell line cultures can be used. Examples include the hybridoma technique, the trioma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique (see, e.g., Cole (1985) in Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, Inc., pp. 77-96).

Techniques described for the production of single chain antibodies (see, e.g., U.S. Patent No. 4,946,778) can be adapted to produce single chain antibodies to the polypeptides of the invention. Alternatively, transgenic mice may be used to express humanized antibodies to these polypeptides or fragments thereof.

Antibodies generated against the polypeptides of the invention may be used in screening for similar polypeptides (e.g., amylases) from other organisms and samples. In such techniques, polypeptides from the organism are contacted with the antibody and those polypeptides which specifically bind the antibody are detected. Any of the procedures described above may be used to detect antibody binding.

Kits

The invention provides kits comprising the compositions, e.g., nucleic acids, expression cassettes, vectors, cells, transgenic seeds or plants or plant parts, polypeptides (e.g., amylases) and/or antibodies of the invention. The kits also can contain instructional material teaching the methodologies and industrial uses of the invention, as described herein.

Measuring Metabolic Parameters

The methods of the invention provide whole cell evolution, or whole cell engineering, of a cell to develop a new cell strain having a new phenotype, e.g., a new or modified amylase activity, by modifying the genetic composition of the cell. The genetic composition can be modified by addition to the cell of a nucleic acid of the invention. To detect the new phenotype, at least one metabolic parameter of a modified cell is monitored in the cell in a "real time" or "on-line" time frame. In one aspect, a plurality of cells, such as a cell culture, is monitored in "real time" or "on-line." In one aspect, a plurality of metabolic parameters is monitored in "real time" or "on-line." Metabolic parameters can be monitored using the amylases of the invention.

Metabolic flux analysis (MFA) is based on a known biochemistry framework. A linearly independent metabolic matrix is constructed based on the law of mass conservation and on the pseudo-steady state hypothesis (PSSH) on the intracellular metabolites. In practicing the methods of the invention, metabolic networks are established, including the:

- identity of all pathway substrates, products and intermediary metabolites
- identity of all the chemical reactions interconverting the pathway metabolites, the stoichiometry of the pathway reactions,
- identity of all the enzymes catalyzing the reactions, the enzyme reaction kinetics,
- the regulatory interactions between pathway components, e.g. allosteric interactions, enzyme-enzyme interactions etc,
- intracellular compartmentalization of enzymes or any other supramolecular organization of the enzymes, and,
- the presence of any concentration gradients of metabolites, enzymes or effector molecules or diffusion barriers to their movement.

Once the metabolic network for a given strain is built, mathematic presentation by matrix notion can be introduced to estimate the intracellular metabolic

fluxes if the on-line metabolome data is available. Metabolic phenotype relies on the changes of the whole metabolic network within a cell. Metabolic phenotype relies on the change of pathway utilization with respect to environmental conditions, genetic regulation, developmental state and the genotype, etc. In one aspect of the methods of the invention, after the on-line MFA calculation, the dynamic behavior of the cells, their phenotype and other properties are analyzed by investigating the pathway utilization. For example, if the glucose supply is increased and the oxygen decreased during the yeast fermentation, the utilization of respiratory pathways will be reduced and/or stopped, and the utilization of the fermentative pathways will dominate. Control of physiological state of cell cultures will become possible after the pathway analysis. The methods of the invention can help determine how to manipulate the fermentation by determining how to change the substrate supply, temperature, use of inducers, etc. to control the physiological state of cells to move along desirable direction. In practicing the methods of the invention, the MFA results can also be compared with transcriptome and proteome data to design experiments and protocols for metabolic engineering or gene shuffling, etc.

In practicing the methods of the invention, any modified or new phenotype can be conferred and detected, including new or improved characteristics in the cell. Any aspect of metabolism or growth can be monitored.

Monitoring expression of an mRNA transcript

In one aspect of the invention, the engineered phenotype comprises increasing or decreasing the expression of an mRNA transcript (e.g., an amylase message) or generating new (e.g., amylase) transcripts in a cell. This increased or decreased expression can be traced by testing for the presence of an amylase of the invention or by amylase activity assays. mRNA transcripts, or messages, also can be detected and quantified by any method known in the art, including, e.g., Northern blots, quantitative amplification reactions, hybridization to arrays, and the like. Quantitative amplification reactions include, e.g., quantitative PCR, including, e.g., quantitative reverse transcription polymerase chain reaction, or RT-PCR; quantitative real time RT-PCR, or "real-time kinetic RT-PCR" (see, e.g., Kreuzer (2001) Br. J. Haematol. 114:313-318; Xia (2001) Transplantation 72:907-914).

In one aspect of the invention, the engineered phenotype is generated by knocking out expression of a homologous gene. The gene's coding sequence or one or more

transcriptional control elements can be knocked out, e.g., promoters or enhancers. Thus, the expression of a transcript can be completely ablated or only decreased.

In one aspect of the invention, the engineered phenotype comprises increasing the expression of a homologous gene. This can be effected by knocking out of a negative control element, including a transcriptional regulatory element acting in cis- or trans- , or, mutagenizing a positive control element. One or more, or, all the transcripts of a cell can be measured by hybridization of a sample comprising transcripts of the cell, or, nucleic acids representative of or complementary to transcripts of a cell, by hybridization to immobilized nucleic acids on an array.

Monitoring expression of a polypeptides, peptides and amino acids

In one aspect of the invention, the engineered phenotype comprises increasing or decreasing the expression of a polypeptide (e.g., an amylase) or generating new polypeptides in a cell. This increased or decreased expression can be traced by determining the amount of amylase present or by amylase activity assays. Polypeptides, peptides and amino acids also can be detected and quantified by any method known in the art, including, e.g., nuclear magnetic resonance (NMR), spectrophotometry, radiography (protein radiolabeling), electrophoresis, capillary electrophoresis, high performance liquid chromatography (HPLC), thin layer chromatography (TLC), hyperdiffusion chromatography, various immunological methods, e.g. immunoprecipitation, immunodiffusion, immuno-electrophoresis, radioimmunoassays (RIAs), enzyme-linked immunosorbent assays (ELISAs), immuno-fluorescent assays, gel electrophoresis (e.g., SDS-PAGE), staining with antibodies, fluorescent activated cell sorter (FACS), pyrolysis mass spectrometry, Fourier-Transform Infrared Spectrometry, Raman spectrometry, GC-MS, and LC-Electrospray and cap-LC-tandem-electrospray mass spectrometries, and the like. Novel bioactivities can also be screened using methods, or variations thereof, described in U.S. Patent No. 6,057,103. Furthermore, as discussed below in detail, one or more, or, all the polypeptides of a cell can be measured using a protein array.

Industrial Applications

Detergent Compositions

The invention provides detergent compositions comprising one or more polypeptides of the invention, for example, amylases of the invention, such as alpha amylases, glucoamylases, etc., and methods of making and using these compositions.

The invention incorporates all methods of making and using detergent compositions, see, e.g., U.S. Patent No. 6,413,928; 6,399,561; 6,365,561; 6,380,147. The detergent compositions can be a one and two part aqueous composition, a non-aqueous liquid composition, a cast solid, a granular form, a particulate form, a compressed tablet, a gel and/or a paste and a slurry form. The invention also provides methods capable of a rapid removal of gross food soils, films of food residue and other minor food compositions using these detergent compositions. Amylases of the invention can facilitate the removal of starchy stains by means of catalytic hydrolysis of the starch polysaccharide. Amylases of the invention can be used in dishwashing detergents in textile laundering detergents.

The actual active enzyme content depends upon the method of manufacture of a detergent composition and is not critical, assuming the detergent solution has the desired enzymatic activity. In one aspect, the amount of amylase present in the final solution ranges from about 0.001 mg to 0.5 mg per gram of the detergent composition. The particular enzyme chosen for use in the process and products of this invention depends upon the conditions of final utility, including the physical product form, use pH, use temperature, and soil types to be degraded or altered. The enzyme can be chosen to provide optimum activity and stability for any given set of utility conditions. In one aspect, the polypeptides of the present invention are active in the pH ranges of from about 4 to about 12 and in the temperature range of from about 20°C to about 95°C. The detergents of the invention can comprise cationic, semi-polar nonionic or zwitterionic surfactants; or, mixtures thereof.

Amylases of the present invention can be formulated into powdered and liquid detergents having pH between 4.0 and 12.0 at levels of about 0.01 to about 5% (preferably 0.1% to 0.5%) by weight. These detergent compositions can also include other enzymes such as known proteases, cellulases, lipases or endoglycosidases, as well as builders and stabilizers. The addition of amylases of the invention to conventional cleaning compositions does not create any special use limitation. In other words, any temperature and pH suitable for the detergent is also suitable for the present compositions as long as the pH is within the above range, and the temperature is below the described enzyme's denaturing temperature. In addition, the polypeptides of the invention can be used in a cleaning composition without detergents, again either alone or in combination with builders and stabilizers.

The present invention provides cleaning compositions including detergent compositions for cleaning hard surfaces, detergent compositions for cleaning fabrics, dishwashing compositions, oral cleaning compositions, denture cleaning compositions, and contact lens cleaning solutions.

In one aspect, the invention provides a method for washing an object comprising contacting the object with a polypeptide of the invention under conditions sufficient for washing. In one aspect, a polypeptide of the invention (e.g., an alkaline-active amylase) is used in a detergent, i.e., as a detergent additive. The detergent composition of the invention may, for example, be formulated as a hand or machine laundry detergent composition comprising a polypeptide of the invention. Detergent compositions of the invention include laundry and dishwashing (e.g., autodishwashing) solutions and application. A laundry additive suitable for pre-treatment of stained fabrics can comprise a polypeptide of the invention. A fabric softener composition can comprise a polypeptide of the invention. Alternatively, a polypeptide of the invention can be formulated as a detergent composition for use in general household hard surface cleaning operations. In alternative aspects, detergent additives and detergent compositions of the invention may comprise one or more other enzymes such as a protease, a lipase, a cutinase, another amylase, a carbohydrase, a cellulase, a pectinase, a mannanase, an arabinase, a galactanase, a xylanase, an oxidase, e.g., a lactase, and/or a peroxidase. The properties of the enzyme(s) of the invention are chosen to be compatible with the selected detergent (i.e. pH-optimum, compatibility with other enzymatic and non-enzymatic ingredients, etc.) and the enzyme(s) is present in effective amounts. In one aspect, amylase enzymes of the invention are used to remove malodorous materials from fabrics. Various detergent compositions and methods for making them that can be used in practicing the invention are described in, e.g., U.S. Patent Nos. 6,333,301; 6,329,333; 6,326,341; 6,297,038; 6,309,871; 6,204,232; 6,197,070; 5,856,164.

Treating fabrics

The invention provides methods of treating fabrics using one or more polypeptides of the invention. The polypeptides of the invention can be used in any fabric-treating method, which are well known in the art, see, e.g., U.S. Patent No. 6,077,316. For example, in one aspect, the feel and appearance of a fabric is improved by

a method comprising contacting the fabric with an amylase of the invention in a solution. In one aspect, the fabric is treated with the solution under pressure.

In one aspect, the enzymes of the invention are applied during or after the weaving of textiles, or during the desizing stage, or one or more additional fabric processing steps. During the weaving of textiles, the threads are exposed to considerable mechanical strain. Prior to weaving on mechanical looms, warp yarns are often coated with sizing starch or starch derivatives in order to increase their tensile strength and to prevent breaking. The enzymes of the invention can be applied to remove these sizing starch or starch derivatives. After the textiles have been woven, a fabric can proceed to a desizing stage. This can be followed by one or more additional fabric processing steps. Desizing is the act of removing size from textiles. After weaving, the size coating must be removed before further processing the fabric in order to ensure a homogeneous and wash-proof result. The invention provides a method of desizing comprising enzymatic hydrolysis of the size by the action of an enzyme of the invention.

The enzymes of the invention can be used to desize fabrics, including cotton-containing fabrics, as detergent additives, e.g., in aqueous compositions. The invention provides methods for producing a stonewashed look on indigo-dyed denim fabric and garments. For the manufacture of clothes, the fabric can be cut and sewn into clothes or garments, which is afterwards finished. In particular, for the manufacture of denim jeans, different enzymatic finishing methods have been developed. The finishing of denim garment normally is initiated with an enzymatic desizing step, during which garments are subjected to the action of amylolytic enzymes in order to provide softness to the fabric and make the cotton more accessible to the subsequent enzymatic finishing steps. The invention provides methods of finishing denim garments (e.g., a "bio-stoning process"), enzymatic desizing and providing softness to fabrics using the amylases of the invention. The invention provides methods for quickly softening denim garments in a desizing and/or finishing process.

Foods and food processing: liquification of starch

The enzymes of the invention have numerous applications in food processing industry. The amylases of the invention are used in starch to fructose processing. Starch to fructose processing can consist of four steps: liquefaction of

granular starch, saccharification of the liquefied starch into dextrose, purification, and isomerization to fructose.

The invention provides methods of starch liquefaction using the enzymes of the invention. Concentrated suspensions of starch polymer granules are converted into a solution of soluble shorter chain length dextrins of low viscosity. This step is useful for convenient handling with standard equipment and for efficient conversion to glucose or 10^3 other sugars. In one aspect, the granular starch is liquefied by gelatinizing the granules by raising the temperature of the granular starch to over about 72°C. The heating process instantaneously disrupts the insoluble starch granules to produce a water soluble starch solution. The solubilized starch solution can then be liquefied by an amylase of the invention. Thus, the invention provides enzymatic starch liquefaction processes using an amylase of the invention.

Figure 26, Figure 27 and Figure 28 illustrate alternative exemplary starch processes, including starch liquefaction processes, of the invention (using at least one enzyme of the invention). For example, Figure 26 illustrates an exemplary starch liquefaction process of the invention comprising treating a starch slurry (e.g., having about 30% to 35% solids) with steam for primary liquefaction (e.g., at about 105°C for about 5 minutes), input into a flash tank, followed by secondary liquefaction (e.g., at about 90°C to 95°C for about 90 minutes), each or one of these steps involving use of an enzyme of the invention. Figure 27 illustrates another exemplary starch liquefaction process of the invention comprising treating a starch slurry at about between pH 4 to pH 5, e.g., pH 4.5, adjusting the pH, calcium addition, liquefaction at about pH 5 to pH 6, e.g., pH 5.4, at about 95°C using an alpha amylase of the invention, followed by another pH and temperature adjustment for saccharification at about between pH 4 to pH 5, e.g., pH 4.5, at a temperature of between about 60°C to 65°C using a glucoamylase of the invention. Figure 28 illustrates another exemplary starch process of the invention comprising treating a starch slurry at about between pH 4 to pH 5, e.g., pH 4.5, (optional adjusting pH, calcium addition), combined liquefaction-saccharification using an alpha amylase and/or a glucoamylase of the invention at about between pH 4 to pH 5, e.g., pH 4.5, at a temperature of greater than about 90°C, or, greater than about 95°C, followed by another pH and temperature adjustment for saccharification at about between pH 4 to pH 5, e.g., pH 4.5, at a temperature of between about 60°C to 65°C using a glucoamylase of the invention. In one aspect, the combined liquefaction-saccharification of the invention

is a "one-pot" process. In one aspect, the entire process is a "one-pot" process. Any one of these processes, and any one of these steps, can also comprise, or can further comprise, another enzyme of the invention (e.g., a glucosidase such as an a-1,6-glucosidase, a maltase, etc.), or another enzyme such as a pullulanase or an isomerase.

An exemplary enzymatic liquefaction process involves adjusting the pH of a granular starch slurry to between 6.0 and 6.5 and the addition of calcium hydroxide, sodium hydroxide or sodium carbonate. In one aspect, calcium hydroxide is added. This provides calcium ions to stabilize the glucoamylase of the invention against inactivation. In one aspect, upon addition of amylase, the suspension is pumped through a steam jet to instantaneously raise the temperature to between 80°-115°C. In one aspect, the starch is immediately gelatinized and, due to the presence of amylase, depolymerized through random hydrolysis of a-1,4-glycosidic bonds by amylase to a fluid mass. The fluid mass can be easily pumped.

The invention provides various enzymatic starch liquefaction processes using an amylase of the invention. In one aspect of the liquefaction process of the invention, an amylase is added to the starch suspension and the suspension is held at a temperature of between about 80°-100°C to partially hydrolyze the starch granules. In one aspect, the partially hydrolyzed starch suspension is pumped through a jet at temperatures in excess of about 105°C to thoroughly gelatinize any remaining granular structure. In one aspect, after cooling the gelatinized starch, a second addition of amylase is made to further hydrolyze the starch.

The invention provides enzymes and processes for hydrolyzing liquid (liquefied) and granular starch. Such starch can be derived from any source, e.g., corn, wheat, milo, sorghum, rye or bulgher. The invention applies to any grain starch source which is useful in liquefaction, e.g., any other grain or vegetable source known to produce starch suitable for liquefaction. The methods of the invention comprise liquefying starch from any natural material, such as rice, germinated rice, corn, barley, milo, wheat, legumes and sweet potato. The liquefying process can substantially hydrolyze the starch to produce a syrup. The temperature range of the liquefaction can be any liquefaction temperature which is known to be effective in liquefying starch. For example, the temperature of the starch can be between about 80°C to about 115°C, between about 100°C to about 110°C, and from about 105°C to about 108°C. In alternative aspects, the amylase used in these methods is active at these temperatures, e.g., active at temperatures

in a range of between about 80°C to about 115°C, between about 100°C to about 110°C, and from about 105°C to about 108°C.

The invention provides methods for liquefaction saccharification as illustrated in Figure 17. In one aspect, alpha-amylases of the invention are used in the illustrated liquefaction step (some current industrial methods use *B. licheniformis* a-amylase). In one aspect, the process takes place at about pH 6.0 at a temperature anywhere in the range of between about 95°C to 105°C, for a length of time anywhere between about 0.5 and 5 hours, e.g., 60, 90 or 120 minutes. In one aspect, in a corn steep process, prior to liquefaction cellulases, proteases and/or protein thioreductases are added.

In one aspect of a liquefaction process of the invention, an amylase of the invention that has activity at about pH 4.5 (or, anywhere between about pH 5 and pH 5), that may or may not be Ca^{2+} dependent is added. Eliminating the addition of salts in the front end of the process eliminates the need to remove them at the back end of the process. In one aspect of a liquefaction process of the invention, an amylase that is more active is used. This can allow one to decrease the amount of enzyme needed. In one aspect, liquefaction and saccharification are done in the same pot, as a "one-pot process," for example, under conditions comprising about 90°C to 95°C (or, anywhere between about 80°C to 105°C), as about a 3 hour process (or, as a process lasting between about 1 and 5 hours). In this aspect, the enzyme load can be cut in half again.

In one aspect of a saccharification process of the invention, a glucoamylase of the invention is used. In one aspect, glucoamylases of the invention are used in the illustrated saccharification step (some current industrial methods use *A. niger* glucoamylase). In one aspect, the process takes place at about pH 4.5, in a temperature range of between about 60°C to 62°C (or, anywhere in the range of between about 50°C to 72°C, or, between about 40°C to 80°C) as a process lasting between about 12 and 96 or more hours. In one aspect of a saccharification process of the invention, a glucoamylase of the invention is used to give a higher level of dextrose in the syrup. In one aspect, other enzymes are added, e.g., pullulanases to increase the amount of glucose.

In one aspect, amylases of the invention are used in the illustrated isomerization step (some current industrial methods use *Streptomyces sp.* glucose isomerase). In one aspect, the isomerization reaction of the invention takes place under conditions comprising anywhere between about pH 5 and pH 10, or anywhere between about ph 6 and pH 9, or anywhere between about pH 7.0 and 8.5. In one aspect, the

isomerization reaction of the invention takes place under conditions comprising between about 40°C to 75°C, or between about 50°C to 65°C, or between about 55°C to 60°C.

In one aspect of an isomerization process of the invention, a xylose isomerase is used. In one aspect, cobalt is used in the reaction (some known thermostable glucose isomerases require cobalt). In one aspect, an enzyme of the invention is used that lacks dependency, or has less dependency, on cobalt. In one aspect, an enzyme of the invention is used that has activity at a lower pH, e.g., pH 7.0, pH 6.5, pH 6, pH 5.5, pH 5, pH 4.5, pH 4, pH 3.5 or less, or, e.g., between a range of about pH 3.5 to 7.0). In one aspect, this allows less color formation (otherwise, excess color may have to be removed). In one aspect, the temperature is increased during isomerization, e.g. to between about 80°C to 110°C, 85°C to 105°C, or 90°C to 100°C. This can increase the amount of fructose produced, e.g. to about 51%. However, in one aspect, for sodas (e.g., soft drinks and the like), the fructose level can be anywhere between about 45% and 65%, or 50% and 60%, e.g., about 55%.

In one aspect, one, some or all of the enzymes used in processes of the invention (including the enzymes of the invention) are immobilized, e.g., immobilized on any surface, e.g., a flat surface or an enzyme column, e.g., immobilized on an array, a bead, fiber, pore, capillary and the like. In one aspect, by being immobilized, they can be reused.

In one aspect, the invention provides "enzyme cocktails" using at least one enzyme of the invention. In one aspect, "enzyme cocktails" are used in the processes of the invention, e.g., including the liquefaction saccharification methods as illustrated in Figure 17. For example, in one aspect, cell wall degrading enzymes (CWDE) are used, e.g., for textile, pulp and paper, and laundry processes of the invention, including, e.g., combinations of cellulases, hemicellulases, xylanase, galactomannanases, glucomannanases, arabinofuranosidases, and others. In one aspect, "enzyme cocktails" used in the processes of the invention for bio-bleaching (e.g., pulp and paper, laundry processes), include combinations of laccases, peroxidases, oxidases and the like. In one aspect, cell wall degrading enzymes are combined with bio-bleaching enzymes and enzymes of the invention to degrade plant cell walls to release color agents.

Processes to produce high MW dextrose syrups

The invention provides processes to produce high MW dextrose syrups using enzymes of the invention, including methods for producing oligosaccharides having a MW tightly groups at about 20,000 MW. In one aspect, amylases of the invention of archael origin, including the archael-derived amylases of SEQ ID NO:80 (encoded by SEQ ID NO:79), SEQ ID NO:82 (encoded by SEQ ID NO:81), SEQ ID NO:116 (encoded by SEQ ID NO:115), SEQ ID NO:323 (encoded by SEQ ID NO:322), SEQ NO: 570 (encoded by SEQ ID NO:169) and enzymes of the invention having the same activity as these archael amylases, are used to liquefy a starch-comprising composition, e.g., a corn starch, to produce an oligosaccharide pattern that is tightly grouped at about 20,000 MW (*Bacillus* amylases will produce syrups containing much higher MW fragments, and high MW oligosaccharides are not fully converted to glucose by glucoamylases, e.g., *Aspergillus* glucoamylases, during saccharification).

Using the amylases of the invention of archael origin to catalyze the hydrolysis of a starch-comprising composition, e.g., a corn starch, the approximately 20,000 MW fragments are produced. These approximately 20,000 MW fragments can be rapidly and fully converted to glucose. Thus, in one aspect, saccharified syrups resulting from *Bacillus* amylase liquefaction contain less dextrose than saccharified syrups from liquefaction using amylases of the invention.

Processes to produce homogenous maltodextrins

The invention provides processes to produce homogenous maltodextrins using enzymes of the invention. The homogenous maltodextrins produced by the methods of the invention can be used in a wide variety of food, drug and coating applications. In one aspect, amylases of the invention of archael origin, including the archael amylases of SEQ ID NO:79, SEQ ID NO:80, SEQ ID NO:81, SEQ ID NO:82, SEQ ID NO:115, SEQ ID NO:116, SEQ ID NO:322, SEQ ID NO:323, and enzymes of the invention having the same activity as these archael amylases, can generate an extremely uniform maltodextrin composition (conventional manufacturing processes using either acid or enzymatic hydrolysis of starch result in a broad, typically bimodal MW distribution of oligosaccharides). The homogenous maltodextrins produced by the methods of the invention have a homogenous MW distribution and can be used in a

variety of maltodextrin-comprising products, resulting in lower viscosity, clear (no haze) solutions, better coating properties, better film-forming properties, and the like.

In one aspect, amylases of the invention of archael origin (and enzymes of the invention having the same activity as these archael amylases) are used to liquefy corn starch to produce a uniform maltodextrin-comprising composition. In one aspect, the liquefication is conducted at a pH of between about pH 4.5 to about pH 6.5, e.g., pH 5.0 or 5.5, at temperatures up to about 105°C. The uniform maltodextrin composition can be produced at DE's ranging from about 5 to as high as about 20. The syrups produced by these archael-derived amylases of the invention can be filtered, treated with charcoal and/or spray-dried to yield the maltodextrin-comprising product.

Enzymatic dry milling processes

The invention provides enzymatic dry milling processes using an amylase of the invention. In dry milling, whole grain is ground and combined with water. The germ is optionally removed by flotation separation or equivalent techniques. The resulting mixture, which contains starch, fiber, protein and other components of the grain, is liquefied using amylase. In one aspect, enzymatic liquefaction is done at lower temperatures than the starch liquification processes discussed above. In one aspect, after gelatinization the starch solution is held at an elevated temperature in the presence of amylase until a DE of 10-20 is achieved. In one aspect, this is a period of about 1-3 hours. Dextrose equivalent (DE) is the industry standard for measuring the concentration of total reducing sugars, calculated as D-glucose on a dry weight basis. Unhydrolyzed granular starch has a DE of virtually zero, whereas the DE of D-glucose is defined as 100.

Enzymatic wet milling processes

The invention provides wet milling processes, e.g., corn wet milling, using an enzyme, e.g., an amylase, of the invention. Corn wet milling is a process which produces corn oil, gluten meal, gluten feed and starch. Thus, the invention provides methods of making corn oil, gluten meal, gluten feed and starch using an enzyme of the invention. In one aspect, an alkaline-amylase of the invention is used in the liquefaction of starch. In one aspect, a glucoamylase of the invention is used in saccharification to produce glucose. An exemplary corn wet milling process of the invention (using at least one enzyme of the invention) is illustrated in Figure 25. Figure 25 illustrates an exemplary corn oil process of the invention comprising steeping, de-germing, de-fibering

and gluten separation, followed by liquefaction using an enzyme of the invention (e.g., an alpha amylase), and saccharification using an enzyme of the invention (e.g., glucoamylase).

In one aspect, corn (a kernel that consists of a outer seed coat (fiber), starch, a combination of starch and glucose and the inner germ), is subjected to a four step process, which results in the production of starch. In one aspect, the corn is steeped, de-germed, de-fibered, and the gluten is separated. In a steeping process the solubles are taken out. The product remaining after removal of the solubles is de-germed, resulting in production of corn oil and production of an oil cake, which is added to the solubles from the steeping step. The remaining product is de-fibered and the fiber solids are added to the oil cake/solubles mixture. This mixture of fiber solids, oil cake and solubles forms a gluten feed. After de-fibering, the remaining product is subjected to gluten separation. This separation results in a gluten meal and starch. The starch is then subjected to liquefaction and saccharification using polypeptides of the invention to produce glucose.

Figure 25 illustrates an exemplary corn wet milling process of the invention (using at least one enzyme of the invention). Figure 26, Figure 27 and Figure 28 illustrate alternative exemplary starch processes, including starch liquefaction processes, of the invention (using at least one enzyme of the invention).

Anti-staling processes

The invention provides anti-staling processes (e.g., of baked products such as bread) using an amylase of the invention. The invention provides methods to slow the increase of the firmness of the crumb (of the baked product) and a decrease of the elasticity of the crumb using an amylase of the invention. Staling of baked products (such as bread) is more serious as time passes between the moment of preparation of the bread product and the moment of consumption. The term staling is used to describe changes undesirable to the consumer in the properties of the bread product after leaving the oven, such as an increase of the firmness of the crumb, a decrease of the elasticity of the crumb, and changes in the crust, which becomes tough and leathery. The firmness of the bread crumb increases further during storage up to a level, which is considered as negative. Amylases of the invention are used to retard staling of the bread as described e.g., in U.S. Patent Nos. 6,197,352; 2,615,810 and 3,026,205; Silberstein (1964) Baker's Digest 38:66-72.

In one aspect, an enzyme of the invention is used to retard the staling of baked products while not hydrolyzing starch into the branched dextrins. Branched dextrins are formed by cleaving off the branched chains of the dextrins generated by a-amylase hydrolysis which cannot be degraded further by the a-amylase. This can produce a gummy crumb in the resulting bread. Accordingly, the invention provides a process for retarding the staling of baked products (e.g., leavened baked products) comprising adding an enzyme of the invention comprising exoamylase activity to a flour or a dough used for producing a baked product. Exoamylases of the invention can have glucoamylase, ß-amylase (which releases maltose in the beta-configuration) and/or maltogenic amylase activity.

The invention also provides a process for preparing a dough or a baked product prepared from the dough which comprises adding an amylase of the invention to the dough in an amount which is effective to retard the staling of the bread. The invention also provides a dough comprising said amylase and a premix comprising flour together with said amylase. Finally, the invention provides an enzymatic baking additive, which contains said amylase.

The invention also provides a high yield process for producing high quality corn fiber gum by treatment of corn fiber with an enzyme of the invention followed by hydrogen peroxide treatment to obtain an extract of milled corn fiber. See, e.g., U.S. Patent No. 6,147,206.

Animal feeds and additives

The invention provides methods for treating animal feeds and additives using amylase enzymes of the invention. The invention provides animal feeds and additives comprising amylases of the invention. In one aspect, treating animal feeds and additives using amylase enzymes of the invention can help in the availability of starch in the animal feed or additive. This can result in release of readily digestible and easily absorbed sugars.

Use of an amylase of the invention can increase the digestive capacity of animals and birds. Use of an amylase of the invention can ensure availability of an adequate nutrient supply for better growth and performance. In one aspect, the enzymes of the invention can be added as feed additives for animals. In another aspect, the animal feed can be treated with amylases prior to animal consumption. In another aspect, the

amylases may be supplied by expressing the enzymes directly in transgenic feed crops (as, e.g., transgenic plants, seeds and the like), such as corn. As discussed above, the invention provides transgenic plants, plant parts and plant cells comprising a nucleic acid sequence encoding a polypeptide of the invention. In one aspect, the nucleic acid is expressed such that the amylase is produced in recoverable quantities. The amylase can be recovered from any plant or plant part. Alternatively, the plant or plant part containing the recombinant polypeptide can be used as such for improving the quality of a food or feed, e.g., improving nutritional value, palatability, and rheological properties, or to destroy an antinutritive factor.

Paper or pulp treatment

The enzymes of the invention can be in paper or pulp treatment or paper deinking. For example, in one aspect, the invention provides a paper treatment process using amylases of the invention. In one aspect, the enzymes of the invention can be used to modify starch in the paper thereby converting it into a liquefied form. In another aspect, paper components of recycled photocopied paper during chemical and enzymatic deinking processes. In one aspect, amylases of the invention can be used in combination with cellulases. The paper can be treated by the following three processes: 1) disintegration in the presence of an enzyme of the invention, 2) disintegration with a deinking chemical and an enzyme of the invention, and/or 3) disintegration after soaking with an enzyme of the invention. The recycled paper treated with amylase can have a higher brightness due to removal of toner particles as compared to the paper treated with just cellulase. While the invention is not limited by any particular mechanism, the effect of an amylase of the invention may be due to its behavior as surface-active agents in pulp suspension.

The invention provides methods of treating paper and paper pulp using one or more polypeptides of the invention. The polypeptides of the invention can be used in any paper- or pulp-treating method, which are well known in the art, see, e.g., U.S. Patent No. 6,241,849; 6,066,233; 5,582,681. For example, in one aspect, the invention provides a method for deinking and decolorizing a printed paper containing a dye, comprising pulping a printed paper to obtain a pulp slurry, and dislodging an ink from the pulp slurry in the presence of an enzyme of the invention (other enzymes can also be added). In another aspect, the invention provides a method for enhancing the freeness of pulp, e.g.,

pulp made from secondary fiber, by adding an enzymatic mixture comprising an enzyme of the invention (can also include other enzymes, e.g., pectinase enzymes) to the pulp and treating under conditions to cause a reaction to produce an enzymatically treated pulp. The freeness of the enzymatically treated pulp is increased from the initial freeness of the secondary fiber pulp without a loss in brightness.

Repulping: treatment of lignocellulosic materials

The invention also provides a method for the treatment of lignocellulosic fibers, wherein the fibers are treated with a polypeptide of the invention, in an amount which is efficient for improving the fiber properties. The amylases of the invention may also be used in the production of lignocellulosic materials such as pulp, paper and cardboard, from starch reinforced waste paper and cardboard, especially where repulping occurs at pH above 7 and where amylases can facilitate the disintegration of the waste material through degradation of the reinforcing starch. The amylases of the invention can be useful in a process for producing a papermaking pulp from starch-coated printed paper. The process may be performed as described in, e.g., WO 95/14807.

An exemplary process comprises disintegrating the paper to produce a pulp, treating with a starch-degrading enzyme before, during or after the disintegrating, and separating ink particles from the pulp after disintegrating and enzyme treatment. See also U.S. Patent No. 6,309,871 and other US patents cited herein. Thus, the invention includes a method for enzymatic deinking of recycled paper pulp, wherein the polypeptide is applied in an amount which is efficient for effective de-inking of the fiber surface.

Waste treatment

The enzymes of the invention can be used in a variety of other industrial applications, e.g., in waste treatment. For example, in one aspect, the invention provides a solid waste digestion process using enzymes of the invention. The methods can comprise reducing the mass and volume of substantially untreated solid waste. Solid waste can be treated with an enzymatic digestive process in the presence of an enzymatic solution (including an enzyme of the invention) at a controlled temperature. This results in a reaction without appreciable bacterial fermentation from added microorganisms. The solid waste is converted into a liquefied waste and any residual solid waste. The resulting

liquefied waste can be separated from said any residual solidified waste. See e.g., U.S. Patent No. 5,709,796.

Oral care products

The invention provides oral care product comprising an amylase of the invention. Exemplary oral care products include toothpastes, dental creams, gels or tooth powders, odontics, mouth washes, pre- or post brushing rinse formulations, chewing gums, lozenges, or candy. See, e.g., U.S. Patent No. 6,264,925.

Brewing and fermenting

The invention provides methods of brewing (e.g., fermenting) beer comprising an amylase of the invention. In one exemplary process, starch-containing raw materials are disintegrated and processed to form a malt. An amylase of the invention is used at any point in the fermentation process. For example, amylases of the invention can be used in the processing of barley malt. The major raw material of beer brewing is barley malt. This can be a three stage process. First, the barley grain can be steeped to increase water content, e.g., to around about 40%. Second, the grain can be germinated by incubation at 15-25°C for 3 to 6 days when enzyme synthesis is stimulated under the control of gibberellins. During this time amylase levels rise significantly. In one aspect, amylases of the invention are added at this (or any other) stage of the process. The action of the amylase results in an increase in fermentable reducing sugars. This can be expressed as the diastatic power, DP, which can rise from around 80 to 190 in 5 days at 12°C.

Amylases of the invention can be used in any beer producing process, as described, e.g., in U.S. Patent No. 5,762,991; 5,536,650; 5,405,624; 5,021,246; 4,788,066.

Use in drilling well and mining operations

The invention also includes methods using enzymes of the invention in well and drilling operations, e.g., gas, oil or other drilling or mining operations. For example, in one aspect, enzymes of the invention are used to increase the flow of production fluids from a subterranean formation, e.g., a well or a mine. In one aspect, the enzymes of the invention are used to remove viscous, starch-containing fluids that can be damaging, e.g., fluids formed during production operations. These starch-containing fluids can be found within a subterranean formation which surrounds a completed well

bore. In one aspect, an amylase of the invention is used in an oil well drilling fluid to aid in the carrying away of drilling mud.

In one aspect, the method comprises allowing production fluids (comprising enzymes of the invention) to flow from the well bore or a mine. The methods can comprise reducing the flow of production fluids from the formation below expected flow rates and formulating an enzyme treatment by blending together an aqueous fluid and a polypeptide of the invention. The methods can comprise pumping the enzyme treatment to a desired location within the well bore or other drilled shaft and allowing the enzyme treatment to degrade the viscous, starch-containing, damaging fluid. The methods can comprise removing the fluid from the subterranean formation to the well or shaft surface. In one aspect, the enzyme treatment is effective to attack the alpha glucosidic linkages in the starch-containing fluid. In one aspect, amylases of the invention are used in mine drilling, well drilling (e.g., gas or oil well drilling), and the like to carry away drilling mud, e.g., while drilling the hole (well bore or shaft).

The enzymes of the invention can be used in any well, shaft or mine drilling operation, many of which are well known in the art. For example, the invention provides methods of introducing an enzyme of the invention, which in one aspect can also comprise an oil or gas field production chemical, into a rock formation comprising oil and/or gas, which comprises passing a microemulsion comprising the enzyme (and, in one aspect, the chemical) down a production well and then into the formation. In one aspect, a production well is subjected to a "shut-in" treatment whereby an aqueous composition comprising an enzyme of the invention is injected into the production well under pressure and "squeezed" into the formation and held there. See, e.g., U.S. Patent No. 6,581,687.

In one aspect, the amylases of the invention used in gas, oil or other drilling or mining operations are active at high or low pH and/or high or low temperatures, e.g., amylases of the invention used in these processes are active under conditions comprising about pH 6.5, pH 6, pH 5.5, pH 5, pH 4.5 or pH 4, or lower, or, under conditions comprising about pH 7, pH 7.5 pH 8.0, pH 8.5, pH 9, pH 9.5, pH 10, pH 10.5 or pH 11 or higher. In one aspect, the amylases of the invention used in these processes are active under conditions comprising a temperature range of anywhere between about 0°C to about 37°C, or, between about 37°C to about 95°C or more, or,

between about 80°C to about 120°C, e.g., 85°C, 90°C, 95°C, 98°C, 100°C, 105°C, 110°C, 115°C, 120°C or more.

Delayed release compositions

The invention provides delayed release or "controlled release" compositions comprising a desired composition coated by a latex polymer, e.g., a latex paint, or equivalent. The delayed release/ controlled release compositions of the invention can comprise any desired composition, including enzymes or any active ingredient, including small molecules, drugs, polysaccharides, lipids, nucleic acids, vitamins, antibiotics, insecticides, and the like. In one aspect, the coating will not readily dissolve at a relatively low temperature but will decompose to release the desired composition (e.g., enzyme) at a relatively higher temperature.

The invention provides methods for the delayed release/ controlled release of compositions wherein the composition is coated by a latex polymer, e.g., a latex paint, or equivalent.

The delayed release/ controlled release compositions and methods of the invention can be used for a variety of medical and industrial applications, for example, in one aspect, delayed release/ controlled release enzyme compositions of the invention comprise enzymes involved in guar fracturing fluids in enhanced oil recovery operations. The oilfield guar degrading application of the invention is facilitated by a coating that will not readily dissolve at low temperature but will decompose to release the enzyme at higher temperatures.

In another aspect, the delayed release/ controlled release enzyme compositions of the invention comprise animal feeds or nutritional supplements comprising, e.g., enzymes, vitamins, antibiotics and/or other food, drug or nutritional supplements. These active compounds in the animal feeds or nutritional supplements are protected from pelleting conditions or gastric digestion by the coating on a delayed release/ controlled release composition of the invention.

In one aspect, the release is a temperature activated release, e.g., the desired composition (e.g., enzyme) is released at an elevated temperature, e.g., between about 37°C to about 95°C or more, e.g., 85°C, 90°C, 95°C, 98°C, 100°C or more. The rate of release can be controlled by the thickness or amount of "barrier" or latex polymer, applied to the desired composition, e.g., a pellet or matrix comprising the desired

composition. Thus, the invention provides pellets or matrices having a range of thicknesses of latex polymer or equivalent and methods of using them.

The invention provides delayed release/ controlled release enzyme compositions, e.g., in one aspect, comprising an enzyme of the invention. In one aspect, the invention provides an enzyme (e.g., an enzyme of the invention), or a pelleted composition comprising an enzyme (e.g., an enzyme of the invention), coated with a latex polymer, e.g., a latex paint, or equivalent. In one aspect, the invention provides methods of making delayed release enzyme compositions comprising coating an enzyme (e.g., an enzyme of the invention), or a pelleted composition comprising an enzyme (e.g., an enzyme of the invention), with a latex polymer, e.g., a latex paint, or equivalent. In one aspect, the invention provides methods of making delayed release/ controlled release compositions comprising coating a desired compound with a latex polymer, e.g., a latex paint, or equivalent.

Latex polymers that are used in the delayed release/ controlled release compositions (e.g., delayed release/ controlled release enzyme compositions) and methods of the invention include, but are not limited to, various types such as the following: acrylics; alkyds; celluloses; coumarone-indenes; epoxys; esters; hydrocarbons; maleics; melamines; natural resins; oleo resins; phenolics; polyamides; polyesters; rosins; silicones; styrenes; terpenes; ureas; urethanes; vinyls; and the like. Latex polymers that are used in the delayed release compositions and methods of the invention also include, but are not limited to, one or more homo- or copolymers containing one or more of the following monomers: (meth)acrylates; vinyl acetate; styrene; ethylene; vinyl chloride; butadiene; vinylidene chloride; vinyl versatate; vinyl propionate; t-butyl acrylate; acrylonitrile; neoprene; maleates; fumarates; and the like, including plasticized or other derivatives thereof.

The amount of latex polymer used in the latex composition of the invention is not critical, but may be any amount following well established procedures using latex polymers. In alternative aspects, the amount of dry latex polymer is at least about 1, or, from about 2 to about 50, or, from about 3 to about 40 weight percent of the total latex composition. The latex composition of the invention may optionally contain other components such as those generally used in latex compositions. These additional components include, but are not limited to, one or more of the following: solvents such as aliphatic or aromatic hydrocarbons, alcohols, esters, ketones, glycols, glycol ethers,

nitroparaffins or the like; pigments; fillers, dryers; flatting agents; plasticizers; stabilizers; dispersants; surfactants; viscosifiers including polymeric associative thickeners, polysaccharide-based thickeners and so on; suspension agents; flow control agents; defoamers; anti-skinning agents; preservatives; extenders; filming aids; crosslinkers; surface improvers; corrosion inhibitors; and other ingredients useful in latex compositions. In one aspect, latex compositions of the invention having improved rheology and stability are provided by combining the latex polymer and a polysaccharide with water following established procedures. See, e.g., U.S. Patent Nos. 6,372,901; 5,610,225.

In one aspect, in making a pelleted or matrix-comprising composition of the invention comprising an active composition, e.g., an enzyme (e.g., an enzyme of the invention), coated with a latex polymer, e.g., a latex paint, or equivalent, the active composition (e.g., enzyme) is embedded in the body of the pellet (in one aspect, a majority, or all, of the active composition (e.g., enzyme) is embedded in the pellet. Thus, harsh chemicals, e.g., the latex coating, which may be an inactivator of the desired, active ingredient, can be used to coat the surface of the pellet or matrix. The composition of the coating can be broken down by agents such as heat, acid, base, pressure, enzymes, other chemicals and the like, to have a controlled release of the desired enzymatic activity triggered by the exposure to the coating-degrading agent.

In one aspect, an active composition, e.g., an enzyme (e.g., an enzyme of the invention, or another enzyme, e.g., a mannanase), is dispersed in a corn term meal and/or a corn starch matrix (e.g., as a pellet). This mixture (e.g., pellet) disintegrates within ten minutes in room temperature (e.g., about 22°C) water to release all (100%) of the active composition, e.g., releases all of the enzymatic activity. At higher temperatures, the rate of release increases. This is not an acceptable rate of disintegration for many uses.

However, as a delayed release/ controlled release composition of the invention, i.e., when this mixture is coated with a latex polymer, e.g., a latex paint, or equivalent, the disintegration of the mixture (e.g., pellet, matrix) is delayed. The rate and extent of release can be controlled by the thickness of the coating (barrier) applied to the pellet or matrix. For example, a coated particle will release only 30% of the activity after six hours in 22°C water. At 60°C, 50% of the enzyme is released in 90 minutes. At 80°C, 80% of the enzyme is released during one hour.

The invention will be further described with reference to the following examples; however, it is to be understood that the invention is not limited to such examples.

EXAMPLES

EXAMPLE 1: Identification and Characterization of Thermostable α-Amylases

The following example describes an exemplary method for determining if a polypeptide is within the scope of the invention. This example describes the identification of novel acid amylases of the invention. The screening program was carried out under neutral and low pH conditions. DNA libraries generated from low pH samples were targeted for discovery. This effort afforded the discovery of hundreds of clones having the ability to degrade starch. DNA sequence and bioinformatic analyses classified many of these genes as previously unidentified amylases.

Biochemical studies

Biochemical analysis of the amylase genomic clones showed that many had pH optima of less than pH 6. Lysates of these genomic clones were tested for thermal tolerance by incubation at 70°C, 80°C, 90°C or 100°C for 10 minutes and measurement of residual activity at pH 4.5. Those clones retaining >50% activity after heat treatment at 80°C were chosen for further analysis. These clones were incubated at 90°C for 10 minutes at pH 6.0 and 4.5 and tested for residual activity at pH 4.5 (Figure 5). A number of clones retained >40% of their activity following this treatment. For comparative purposes, residual activity of an enzyme of the invention (an "evolved" amylase), SEQ ID NO:437 (encoded by SEQ ID NO:436), was equivalent to the best of the second-generation enzymes; the specific activity of SEQ ID NO:437 was greater.

Thermal activity of the clones with residual activity after heat treatment at 90°C at pH 4.5 was measured at room temperature, 70°C and 90°C at pH 4.5. Table 1 shows that the hydrolysis rates of SEQ ID NO: 87 (*B. stearothermophilus* amylase) and SEQ ID NO. 113 (*B. licheniformis* amylase) decrease at higher temperatures, whereas the rate for SEQ ID NO:125 continues to increase as the temperature is raised to 70°C and only reduces by around 50% at 90°C.

The exemplary polypeptide having a sequence as set forth in SEQ ID NO:437 (encoded by SEQ ID NO:436) is thermostable, retaining 50% activity after 25 minutes at 100°C in the absence of added calcium, at pH 4.5. This exemplary polypeptide retained 90% activity after 60 minutes at 100°C in the presence of 40 mg/L calcium, pH 4.5. The activity profile of the polypeptide SEQ ID NO:437 is in the range of between about 4.8 and 5.0. Added calcium is not required for activity.

The polypeptide SEQ ID NO:437 can have a light brown to yellow liquid with a specific gravity of 1.1, at pH 10, when formulated with 35% glycerol. Its alpha amylase activity is between about 110 to 115 IAU* / gram (*IAU = INNOVASE™ activity unit). One analytical method used comprised hydrolysis of 4-nitrophenyl-alpha-D-hexa-glucopyranoside (this same method can be used to determine if an enzyme is within the scope of the invention).

Candidate evaluation

Based on residual activity at pH 4.5 after a 90°C heat treatment, specific activity and rate of starch hydrolysis at 90°C when compared with *B. licheniformis* amylase, SEQ ID NO:125 is compared with the enzyme (an "evolved" amylase) of SEQ ID NO:437 in a starch liquefaction assay.

Table 1.	Room temperature	70°C	90°C
SEQ ID NO.:87[1]	1.25	1.43	0.33
SEQ ID NO.: 113[2]	3.3	1.9	0.39
SEQ ID NO.: 125	1.9	47	19

Table 1 shows rates of dye labeled starch hydrolysis (relative fluorescence units/s) of three genomic clones at pH 4.5 and 3 different temperatures. [1]*B. stearothermophilus* amylase, [2]*B. licheniformis* amylase.

The following table is a summary of Average Relative Activity (ARA), Thermal Tolerance, Thermal Stability, Specific Activity and Expression (Units / L) for selected exemplary enzymes of the invention (for example, SEQ ID NOS: 125, 126, refers to a polypeptide having a sequence as set forth in SEQ ID NO:126, encoded by SEQ ID NO:125, etc.):

Enzyme	Expression Host	Average Relative Activity (ARA)	pH Optimum	Thermal Tolerance %RA after 5 min** 50, 60, 70, 80, 90°C	Thermal Stability %RA 37, 65, 80°C	Specific Activity (U/mg at pH 5.3, 37°C)	Expression (Units / L)
Benchmark SEQ ID NOS:		80	4.0 to 5.5	105, 107, 88, 58, 27	100, 83, 0	82	
125, 126	*Pichia*	66	4.5 to 6.0	86, 88, 100, 86, 65	100, 347, 553	81	8521
378, 379	*Pichia*	66	6.0 to 7.0	22, 0, 0, 0, 0		937	183615
416, 417	*Pichia*	59	4.5 to 5.0	56, 1, 1, 0, 1		39	23256
203, 204	*Pichia*	61	6.0 to 7.0	18, 2, 3, 2, 3		20	122107
434, 435	*Pichia*	76	6.0 to 6.5	151, 58, 0, 0, 0		151	17171
420, 421	*Pichia*	84	5.5 to 7.0	68, 26, 0, 0, 0		75	5005
350, 351	*Pichia*	59	6.0 to 7.0	6, 0, 0, 0, 0		104	39662
402, 403	*Pichia*	67	5.5 to 6.0	42, 8, 11, 12, 16		535	75053
336, 337	*Pichia*	63	4.5 to 5.5	124, 105, 115, 108, 117	100, 0, 0	572	20822
430, 431	*Pichia*	50	6.0 to 6.5	111, 86, 82, 89, 35		138	6556
127, 128	*Pichia*	71	5.5 to 6.5	127, 115, 53, 4, 5		17	114999

101, 102	*Pichia*	63	5.0 to 5.5	124, 164, 145, 120, 144		28	11559
388, 389	*Pichia*	80	6.0 to 7.0	87, 29, 5, 0, 0		259	163163
539, 540	*Pichia*	TBD	4.0 to 4.5	102, 100, 31, 12, 3	100, 186, 123	TBD	TBD

A.R.A. is Average Relative Activity. A.R.A. is calculated as the average relative activity of an amylase between pH 4 and pH 7.5.

[#] Approximate units per liter expression is calculated as follows: (total units of amylase present in recovered lyophilized powder) (volume of culture in fermenter)

Evaluation of the amylase SEQ ID NO:437

The amylase SEQ ID NO:437 (encoded by SEQ ID NO:436) was evaluated under a variety of conditions. In the following protocols N°2 yellow dent corn was used as a starch source.

Liquefaction

A starch slurry comprising 35% dry solids ("DS") was subjected to primary liquefaction for five minutes under various temperatures in the range of 95°C to 119°C (e.g., at about 110°C), with an enzyme concentration of between 0.2 to 0.8 gram/kilogram (g/kg) starch DS, with added calcium in the range of between zero and 30 parts per million (ppm), at pH 4.0 to pH 5.6. Secondary liquefaction comprised conditions of 120 minutes at 95°C.

Saccharification

Saccharification was initially tested using 35% dry solids ("DS") (starch slurry) and glucoamylase AMG 300L (Novozymes A/S, Denmark) at 0.225 AGU/gram DS (AGU= amyloglucosidase, or glucoamylase, units), pH 4.3, at 60°C for 44 hours.

The amylase SEQ ID NO:437 was demonstrated to be useful under the above-described pH conditions, was calcium independent and had a high thermal stability. In one aspect, amylase SEQ ID NO:437, or another amylase of the invention, is used in a dosage range of between 0.5 to 0.7 kg / MT DS starch.

The invention provides methods for making nutritive sweeteners using enzymes of the invention, e.g., processes comprising the above described liquefaction and

saccharification protocols using, e.g., amylase SEQ ID NO:437, or another enzyme of the invention. In one aspect, the dosage range for an enzyme of the invention in these processes is between about 0.5 to 0.7 gram per kg starch DS, a jet temperature (e.g., using a jet cooker) of about 110°C, pH 4.5, no added calcium.

Dry Mill Ethanol Production

The invention provides methods for Dry Mill Ethanol Production using enzymes of the invention, e.g., amylase SEQ ID NO:437, or another enzyme of the invention.

In evaluating enzymes of the invention for use in Dry Mill Ethanol Production, particularly, liquefaction of dry mill corn flour, a bench scale reactor was used with corn flour sourced from commercial dry mill. TERMAMYL™ SC (Novozymes A/S, Denmark) amylase was used as a competitive benchmark. Test found optimum conditions to be 85°C, pH 5.7. Five independent variables were studied: temperature (in a range of between 80°C to 100°C), enzyme dose of between 0.2 to 1.0 g/kg starch, pH 4.4 to 6.0, calcium in a range between 0 ppm to 200 ppm, and a recycled backset between about 0% to 40%.

At 95°C amylase SEQ ID NO:437 reduces viscosity of dry mill corn flour more rapidly than TERMAMYL™ SC (Novozymes A/S, Denmark) amylase at its optimum conditions, including at 85°C. The rate of viscosity reduction by amylase SEQ ID NO:437 was influenced most by enzyme dose and temperature. The optimal range was found to be in the range of 0.4 to 0.6 g/kg starch, with an optimum temperature at 95°C. The amylase SEQ ID NO:437 was effective at a lower pH and a higher temperature than TERMAMYL™ SC (Novozymes A/S, Denmark) amylase at a pH in the range between pH 4.4 and pH 5.6. Calcium addition had a minimal effect on rate of viscosity reduction at 95°C. The amylase SEQ ID NO:437 was effective in the presence of a 30% recycled backset (e.g., thin stillage, spent wash = recycling byproducts back into liquefaction). Figure 29 shows data summarizing these findings comparing amylase SEQ ID NO:437 with TERMAMYL™ SC (Novozymes A/S, Denmark) amylase in dry mill ethanol processing.

In alternative aspects, use of amylase SEQ ID NO:437 in dry mill ethanol processes can provide operational advantages, for example: rapid reduction in viscosity of slurried corn flour, making an increase in dissolved solids and throughput possible

without additional capital investment; superior thermal stability to best competitor, which eliminates split dosing (amylase SEQ ID NO:437 is a thermostable enzyme and eliminates the need to dose before jet cooking and after), lower viscosities are obtained at higher process temperatures, and provides improved microbial control in slurry tank (process is run at higher temperature, so unwanted microbes are killed); lower liquefaction pH, which eliminates need for pH adjustment, decreases scale formation (calcium oxalate precipitate forms on hardware, etc.; if liquefaction done at low pH, there is a higher potential for scale formation) and reduces byproduct formation.

In summary, amylase SEQ ID NO:437 is a thermostable enzyme that can meet key industry needs, for example, under certain conditions, rapidly reduces viscosity of high dry solids corn flour slurry, can be thermostable (optimum temperature 95°C), can be calcium independent, can be active under low pH optimum, and can tolerate up to 30% recycled backset. In one aspect, the recommended dose is in the range of between about 0.4 to 0.6 kg/ MT starch.

EXAMPLE 2: Thermostable Amylases Active at Alkaline pH

The following example describes an exemplary method for determining if a polypeptide is within the scope of the invention, e.g., is a thermostable amylase.

The initial focus of this example was the evaluation of an existing panel of amylases in an commercial automatic dish wash (ADW) formulation. This effort identified two candidates: one with activity at high pH (SEQ ID NO.:115) and another with stability in the ADW formulation (SEQ ID NO.:207). Studies also included the identification of high pH amylases. This effort afforded the discovery of hundreds of clones having the ability to degrade starch. DNA sequence and bioinformatics analyses classified many of these genes as previously unidentified amylases. The remaining open reading frames were neopullulanases, amylopullulanases and amylomaltases. Extensive biochemical and applications studies showed that 3 candidates: clone B, SEQ ID NO.:147 and SEQ ID NO.:139) have high specific activity at pH10, but unfortunately lack stability in the ADW formulation. In summary, a panel of novel amylases each having desirable phenotypes for the ADW application has been identified.

Biochemical studies

Biochemical analysis of the amylase genomic clones showed that many of them hydrolyzed starch at pH 10 and 50°C. To produce sufficient quantities of enzyme for further biochemical and applications testing, the amylase open reading frames of the 40 most active genomic clones were subcloned into expression vectors. This effort included making 2 constructs for those clones containing a putative signal sequence and establishing the growth and induction conditions for each subclone (plus and minus the amylase signal peptide).

Soluble, active protein was successfully purified to homogeneity from 34 subclones and specific activity (units/mg, where 1 unit = μmol reducing sugars/min) was measured at pH 8 and pH 10 (40°C and 50°C) using 2% starch in buffer. The amylase from *Bacillus licheniformis* (SEQ ID NO.:113) was chosen as the benchmark for these studies. Specific activity was determined by removing samples at various time points during a 30 minute reaction and analyzing for reducing sugars. The initial rate was determined by fitting the progress curves to a linear equation. A comparison of the top candidates is shown in Table 2.

A study to determine the dependence of hydrolysis rate on pH showed that only clone B is an "alkaline amylase" with a pH optimum of approximately 8; all others had pH optima of 7 or less. Nevertheless, it is clear that the panel of hits included several lead amylases with appreciable activity at pH 10 and 50°C.

Table 2. Specific activities (U/mg pure enzyme) of amylases

Enzyme	Specific activity pH 8, 40°C	Specific activity pH 10, 50°C
Clone B	682	20
SEQ ID NO.:139	430	33
SEQ ID NO.:127	250	47
SEQ ID NO.:137	230	3
SEQ ID NO.:113 (*B. licheniformis*)	228	27
SEQ ID NO.:205	163	4
Remainder	<40	

Stability

Stability in the presence of the ADW formulation was measured for each of the 3 top candidates identified via biochemical analysis. The benchmark for these studies was a commercial enzyme in the formulation matrix. Figure 13 illustrates the

residual activity (measured at pH 8 and 50°C) after a 30 minute incubation at 50°C in the presence of various components of the ADW formulation; pH 8, pH 10.8, ADW solution (with bleach) and ADW solution (without bleach). The measured activity after the incubation is expressed as a percentage of the original activity. The data show that clone B was very sensitive to high temperature, whereas the other amylases were less affected. When the enzymes were incubated at high pH and temperature, the commercial enzyme SEQ ID NO.: 139 became less stable; however, SEQ ID NO.: 127 retained full activity. The apparently anomalous behavior of SEQ ID NO.: 127 after pH 10 incubation vs pH 8 was observed in repeated trials.

When amylase activity on dye-labeled starch is measured in the ADW matrix at 50°C, the commercial amylase exhibits roughly 5% of its activity at pH 8. In the same assay, clone B, SEQ ID NO.: 139 and SEQ ID NO.: 127 exhibit <2% of their original activity measured at pH 8.

Wash tests

Wash tests using starch coated slides were carried out to gauge the performance of each of the purified enzymes as compared to the commercial amylase. The spaghetti starch coated slides were prepared according to protocol. Two pre-weighed starch coated slides were placed back to back in a 50 mL conical tube and 25 mL of ADW solution, +/- enzyme were added per tube. The tubes were incubated for 20 minutes at 50°C with gentle rotation on a vertical carousel. Following the incubation period, the slides were immediately rinsed in water and oven dried overnight. All trials were run in duplicate and the commercial enzyme was run as a positive control. The results (Figure 6) of these experiments are expressed as net % starch removed, e.g. % of starch removed in ADW with enzyme, *minus* the % of starch removed in ADW alone.

EXAMPLE 3: Gene Optimization

The following example describes an exemplary method for determining if a polypeptide is within the scope of the invention, e.g., assessing enzyme performance in the presence of ADW performance.

The properties of enzymes may be improved by various evolution strategies, including GeneSiteSaturationMutagenesis (GSSM™) and GeneReassembly™. (Diversa Corporation, San Diego, CA). Such techniques will be applied to the amylase

nucleic acids of the invention in order to generate pools of variants that can be screened for improved performance. In one aspect, parental molecules for evolution include any nucleic acid of the invention, e.g., are one or all of the following: SEQ ID NO.: 113, SEQ ID NO.: 139, SEQ ID NO.:115 and SEQ ID NO.: 127 (a truncated form of SEQ ID NO.: 127).

A high throughput screen has been developed to assess enzyme performance in the presence of ADW performance. Development of a HTS is of paramount importance in any evolution program The HTS is automated and has showed consistent results for the parental amylases (Figure 7). Parental amylases have measurable activity in the ADW formulation, however highly reduced relative to pH 8 activity.

EXAMPLE 4: Characterization of α-Amylases having Activity at Alkaline pH

The following example describes an exemplary method for determining if a polypeptide is within the scope of the invention, for example, has alpha-amylase activity at alkaline pH.

Amylases of the invention having activity at alkaline pH were characterized further. Kinetics on 2% starch at pH 8 and 10 (40°C and 50°C) have been performed.

Table 4:

Clones, specific activities	pH 8, 40°C	pH 10, 50°C
SEQ ID NO.: 113 (*B. lichenoformis*)	228 units/mg	27 units/mg
Clone B	682 units/mg	31 units/mg
SEQ ID NO.: 139	430 units/mg	33 units/mg
SEQ ID NO.: 127	540 units/mg	50 units/mg
control 0GL5 (*E. coli*)	1.8 units/mg	0 units/mg

1 unit of activity is defined as release of 1 μmol reducing sugars per minute.

EXAMPLE 5: Amylase Activity Assay: BCA Reducing Ends Assay

The following example describes an exemplary method for determining if a polypeptide is within the scope of the invention, for example, by a BCA reducing ends

assay. Amylase activity of clones of interest was determined using the following methodology.

1. Prepare 2 substrate solutions, as follows:

a) 2% soluble starch (potato) pH 8 solution by dissolving 2 gm potato starch in 100 ml 100 mM sodium phosphate pH 8).

b) 2% soluble starch (potato) pH 10 solution by dissolving 2 gm potato starch in 100 ml 100 mM sodium carbonate.

Heat both solutions in a boiling water bath, while mixing, for 30-40 minutes until starch dissolves.

2. Prepare Solution A from 64 mg/ml sodium carbonate monohydrate, 24 mg/ml sodium bicarbonate and 1.95 mg/ml BCA (4,4'-dicarboxy-2,2'- biquinoline disodium salt (Sigma Chemical cat # D-8284). Added above to dH2O.

3. Prepare solution B by combining 1.24 mg/ml cupric sulfate pentahydrate and 1.26 mg/ml L-serine. Add mixture to dH2O.

4. Prepare a working reagent of a 1:1 ration of solutions A and B.

5. Prepare a Maltose standard solution of 10 mM Maltose in dH2O, where the 10 mM maltose is combined in 2% soluble starch at desired pH to a final concentration of 0, 100, 200, 300, 400, 600 µM. The standard curve will be generated for each set of time-points. Since the curve is determined by adding 10 ul of the standards to the working reagent it works out to 0, 1, 2, 3, 4, 6 nmole maltose.

6. Aliquot 1 ml of substrate solution into microcentrifuge tubes, equilibrate to desired temperature (5 min) in heat block or heated water bath. Add 50 ul of enzyme solution to the inside of the tube lid.

7. While solution is equilibrating mix 5 ml of both solution A & B. Aliquot 100 ul to 96 well PCR plate. Set plate on ice.

8. After 5 minute temperature equilibration, close lid on tubes, invert and vortex 3 times. Immediately aliquot 10 ul into plate as t=0 (zero time point). Leave enzyme mixture in heat block and aliquot 10 ul at each desired time point (e.g. 0, 5, 10,15, 20, 30 minutes).

9. Ensure that 12 wells are left empty (only working reagent aliquotted) for the addition of 10 ul of standards, for the standard curve.

10. When all time points are collected and standards are added, cover plate and heated to 80° C for 35 min. Cool plate on ice for 10 min. Add 100 ul H2O to

all wells. Mix and aliquot 100 ul into flat bottomed 96-well plate and read absorbance at 560 nm.

11. Zero each sample's time points against its own t=0 (subtract the average t=0 A560 value from other average A560 values). Convert the $A560_{(experimental)}$ to umole (Divide $A560_{(experimental)}$ by the slope of the standard curve (A560/umole). Generate a slope of the time points and the umole (in umole/min), multiply by 100 (as the umole value only accounts for the 10 ul used in the assay, not the amount made in the 1ml rxn). To get the specific activity divide the slope (in umole/min) by the mg of protein. All points should be done at a minimum in duplicate with three being best. An example standard curve is set forth in Figure 11.

Table 5: Sample data:

Clone	Dilution	Minutes	A560-1	A560-2	Avg A 560	Zeroed A 560	(A560exp/std slope) umole
ENZ	50	0	0.1711	0.1736	0.17235	0	0.0000
		5	0.2104	0.2165	0.21345	0.0411	0.0005
		10	0.2492	0.2481	0.24865	0.0763	0.0009
		15	0.2984	0.2882	0.2933	0.12095	0.0014
		20	0.3355	0.3409	0.3382	0.16585	0.0020
		30	0.3942	0.3805	0.38735	0.215	0.0026
		40	0.4501	0.4412	0.44565	0.2733	0.0033

Activity =0.008646 umole/min

Divide protein concentration (mg/ml) by any dilution to get mg used in assay.

Divide the above slope by mg used in assay to get specific activity

Specific Activity =24.93 umole/min/mg

See for example, Dominic W.S. Wong, Sarah B. Batt, and George H. Robertson (2000) J. Agric. Food Chem. 48:4540-4543; Jeffrey D. Fox and John F. Robyt, (1991) Anal. Biochem. 195, 93-96.

EXAMPLE 6: Screening for α-Amylase activity

The following example describes an exemplary method for determining if a polypeptide is within the scope of the invention. Amylase activity of clones can be assessed by a number of methods known in the art. The following is the general methodology that was used in the present invention. The number of plaques screened, per plate, should be approximately 10,000 pfu's. For each DNA library: at least 50,000

plaques per isolated library and 200,000 plaques per non-isolated library should be screened depending upon the pfu titer for the λ Zap Express amplified lysate.

Titer determination of Lambda Library

1) µL of Lambda Zap Express amplified library stock added to 600µL *E. coli* MRF' cells (OD_{600}=1.0). To dilute MRF' stock, 10mM $MgSO_4$ is used.
2) Incubate at 37 °C for 15 minutes.
3) Transfer suspension to 5-6mL of NZY top agar at 50 °C and gently mix.
4) Immediately pour agar solution onto large (150mm) NZY media plate.
5) Allow top agar to solidify completely (approximately 30 minutes), then invert plate.
6) Incubate the plate at 39 °C for 8-12 hours.
7) Number of plaques is approximated. Phage titer determined to give 10,000 pfu/plate. Dilute an aliquot of Library phage with SM buffer if needed.

Substrate screening

1) Lambda Zap Express (50,000 pfu) from amplified library added to 600µL of *E. coli* MRF' cells (OD600=1.0). For non-environment libraries, prepare 4 tubes (50,000 pfu per tube).
2) Incubate at 37 °C for 15 minutes.
3) While phage/cell suspensions are incubating, 1.0mL of red starch substrate (1.2% w/v) is added to 6.0mL NZY top agar at 50 °C and mixed thoroughly. Keep solution at 50°C until needed.
4) Transfer 1/5 (10,000 pfu) of the cell suspension to substrate/top agar solution and gently mixed.
5) Solution is immediately poured onto large (150mm) NZY media plate.
6) Allow top agar to solidify completely (approximately 30 minutes), then invert plate.
7) Repeat procedures 4-6 4 times for the rest of the cell suspension (1/5 of the suspension each time).
8) Incubate plates at 39 °C for 8-12 hours.
9) Plate observed for clearing zones (halos) around plaques.
10) Plaques with halos are cored out of agar and transferred to a sterile micro tube. A large bore 200µL pipette tip works well to remove (core) the agar plug containing the desired plaque.

11) Phages are re-suspended in 500µL SM buffer. 20µL Chloroform is added to inhibit any further cell growth.

12) Pure phage suspension is incubated at room temperature for 4 hours or overnight before next step.

Isolation of pure clones

1) 10µL of re-suspended phage suspension is added to 500µL of *E. coli* MRF' cells (OD600=1.0).
2) Incubate at 37 °C for 15 minutes.
3) While phage/cell suspension is incubating, 1mL of red starch substrate (1.2% w/v) is added to 6.0mL NZY top agar at 50 °C and mixed thoroughly. Keep solution at 50 °C until needed.
4) Cell suspension is transferred to substrate/top agar solution and gently mixed.
5) Solution is immediately poured onto large (150mm) NZY media plate.
6) Allow top agar to solidify completely (approximately 30 minutes), then invert plate.
7) Plate incubated at 39 °C for 8-12 hours.
8) Plate observed for a clearing zone (halo) around a single plaque (pure clone). If a single plaque cannot be isolated, adjust titer and re-plate phage suspension.
9) Single plaque with halo is cored out of agar and transferred to a sterile micro tube. A large bore 200µL pipette tip works well to remove (core) the agar plug containing the desired plaque. To amplify the titer, core 5 single active plaques into a micro tube.
10) Phages are re-suspended in 500µL SM buffer. 20µL Chloroform is added to inhibit any further cell growth.
11) Pure phage suspension is incubated at room temperature for 4 hours or overnight before next step. The pure phage suspension is stored at –80 °C by adding DMSO into the phage suspension (7% v/v).

Excision of pure clone

1) 100µL of pure phage suspension is added to 200µL *E. coli* MRF' cells (OD600=1.0). To this, 1.0µL of EXASSIST helper phage (>1 x 106 pfu/mL; Stratagene) is added. Use 2059 Falcon tube for excision.
2) Suspension is incubated at 37 °C for 15 minutes.
3) 3.0 mL of 2 x YT media is added to cell suspension.
4) Incubate at 30 °C for at least 6 hours or overnight while shaking.

5) Tube transferred to 70 °C for 20 minutes. The phagemid suspension can be stored at 4 °C for 1 to 2 months.

6) 100 µL of phagemid suspension transferred to a micro tube containing 200µL of *E. coli* Exp 505 cells (OD600=1.0).

7) Suspension incubated at 37 °C for 15 minutes.

8) 300µL of SOB is added to the suspension.

9) Suspension is incubated at 37 °C for 30 to 45 minutes.

10) 100µL of suspension is transferred to a small (90mm) LB media plate containing Kanamycin (LB media with Kanamycin 50µg/mL) for Zap Express DNA libraries or Ampicillin (LB media with Kanamycin 100µg/mL) for Zap II DNA libraries.

11) The rest of suspension is transferred to another small LB media plate.

12) Use sterile glass beads to evenly distribute suspension on the plate.

13) Plates are incubated at 30 °C for 12 to 24 hours.

14) Plate observed for colonies.

15) 15) Inoculate single colony into LB liquid media containing suitable antibiotic and incubate at 30 °C for 12 to 24 hours.

16) Glycerol stock can be prepared by adding 80% glycerol into liquid culture (15% v/v) and stored at –80 °C.

Activity verification

1) 50µL of liquid culture is transferred to a micro tube. Add 500µL of 8% pH7 Amylopectin Azure into the same tube. Prepare 2 tubes for each clone.

2) Activity is tested at 50 °C for 3 hours and overnight. Use pH 7 buffer as control.

3) Cool the test specimen at ice-water bath for 5 minutes.

4) Add 750µL of Ethaqnol and mixed thoroughly.

5) Centrifuge at 13000 rpm (16000 g's) for 5 minutes.

6) Measure OD of the supernatant at 595nm.

RFLP analysis

1) 1.0mL of liquid culture is transferred to a sterile micro tube.

2) Centrifuge at 13200 rpm (16000 g's) for 1 minute.

3) Discard the supernatant. Add another 1.0 mL of liquid culture into the same sterile micro tube.

4) Centrifuge at 13200 rpm (16000 g's) for 1 minute.

5) Discard the supernatant.

6) Follow QIAprep spin mini kit protocol for plasmid isolation.

7) Check DNA concentration using BioPhotometer.

8) Use Sac I and Kpn I for first double digestion. Incubate at 37 °C for 1 hour.

9) Use Pst I and Xho I for second double digestion. Incubate at 37 °C for 1 hour.

10) Add Loading dye into the digested sample.

11) Run the digested sample on a 1.0% agarose gel for 1-1.5 hours at 120 volts.

12) View gel with gel imager. All clones with a different digest pattern will be sent for sequence analysis.

EXAMPLE 7: Assay for Amylases

The following example describes an exemplary method for determining if a polypeptide is within the scope of the invention.

Preparation Of Host Cultures

1. Start an overnight culture of XL1-Blue MRF' host cells. Use a single colony from a streak plate to inoculate 10 mL LB supplemented with 20 ug/mL tetracycline. Grow overnight culture shaking at 37°C for at least 16 hours.
2. Using aseptic technique, inoculate a fresh 100 mL of LB_{Tet} day culture with XL1-Blue MRF' host from the overnight LB_{Tet} culture.
3. Grow in a 37°C shaker until the OD reaches 0.75 – 1.0.
4. Pellet host cells at 1000 x g for 10 minutes and gently resuspend in 10 mM $MgSO_4$ at OD5.
5. Dilute a small amount of host cells to OD1 for use in titering and pintooling.
6. Host preparations can be used for up to 1 week when stored on ice or at 4°C.

-To shorten growth time for the day culture, use ½X the usual Tet concentration in LB (½X = 10 ug/mL), or omit the antibiotic altogether.

-Do not use NZY when selecting with Tetracycline. The high Mg^{++} concentration in NZY medium renders Tet inactive.

Titering Lambda Libraries

7. Place three sterile microfuge tubes in a rack.
8. Aliquot 995 uL prepared host cells in one tube and 45 uL prepared OD1 host cells into each of the two remaining tubes.

9. Add 5 uL of lambda library to the tube containing 995 uL host cells and mix by vortexing. This results in a dilution factor of 200.
10. Prepare 1/2,000 and 1/20,000 dilutions by consecutively adding 5 uL of previous dilution to the remaining two tubes containing 45 uL prepared host cells. Mix by vortexing after each dilution was made.
11. Allow phage to adsorb to host by incubating at 37°C for 15 minutes.
12. Meanwhile, pipet 100 uL of prepared OD1 host cells to each of three Falcon 2059 tubes.
13. Add 5 uL of each dilution to a separate 2059 tube containing host cells.
14. Plate each by adding 3 mL top agar to each tube and quickly pour over 90 mm NZY plates. Ensure a smooth, even distribution before the top agar hardens.
15. Invert plates and incubate at 37°C overnight.
16. Count plaques and calculate titer of the library stock (in plaque forming units (pfu) per uL).

Lambda Microtiter Screening For Amylases

Preparation

1. Prepare a sufficient amount of XL1-Blue MRF' host culture, as described above, for the amount of screening planned. A culture of 100 mL is usually sufficient for screening 2-3 libraries.
2. Autoclave several bottles compatible with the QFill2 dispenser. These are the wide-mouth Corning bottles, 250 mL containing a sealing ring around the lip.
3. Make sure there are sufficient amounts of plates, top agar, BODIPY starch, red starch solution, etc. available for the screen.
4. Schedule the Day 2 robot run with a representative from Automation.

Day 1

1. Label the 1536-well plates (black) with library screen and plate number. Tough-Tags™ tube stickers, cut in half width-wise, are ideal for labeling 1536 well plates.
2. Calculate volumes of library, host cells and NZY medium necessary for the screen. This is easily done with an Excel spreadsheet.
3. Combine the calculated volumes of lambda library and OD5 host cells in a sterile 250 mL wide-mouth Corning bottle (containing a sealing ring).
4. Allow adsorption to occur at 37°C for 15 minutes.

5. Add the calculated volume of NZY medium and mix well. This is referred to as the cell-phage-medium suspension.

6. Perform a concomitant titer by combining 50 uL of the cell-phage-medium suspension with 250 uL of OD1 host cells in a Falcon 2059 tube, then plating with 9 mL of top agar onto a 150 mm NZY plate. Incubate concomitant titer plate at 37°C overnight.

7. Load the dispenser with the remainder of the suspension and array each labeled 1536-well plate at 4 uL per well. If the dispenser leaves air bubbles in some wells, they can be removed by centrifuging the plates at 200 x g for 1 minute.

8. Add 0.5 uL of positive control phage to well position AD46 of at least two of the assay plates. Use a strong amylase-positive lambda clone for this purpose. The lambda versions of SEQ ID NO.: 113 or SEQ ID NO.: 199 are good choices for positive controls.

9. Incubate assay plates at 37°C overnight in a humidified (≥95%) incubator.

Day 2

1. Count the pfu on the concomitant titer plate and determine the average seed density per well (in pfu per well).
2. Pintool at least 2 plates of each library screen (preferably the 2 containing positive controls) as follows:
   a) Prepare 2 host lawn plates to act as a surface on which to pintool: combine 250 uL of OD1 host cells with 2 mL 2% red starch and plate with 9 mL top agar onto 150 mm NZY plates. Hold each plate as level as possible as the top agar solidifies in order to produce an even hue of red across the plate.
   b) Using a twice flame-sterilized 1536 position pintool, replicate at least 2 of the screening plates onto the host lawn plates.
   c) Place the pintooled recipient plates in a laminar flow hood with the lids off for about 15-30 minutes (to vent off excess moisture).
   d) Replace the lids and incubate inverted at 37°C overnight.
3. Prepare the 2X BODIPY starch substrate buffer as follows:
   a) Calculate the total volume of 2X substrate buffer solution needed for all screening plates at 4 uL per well (including any extra deadspace volume

required by the dispenser) and measure this amount of 100 mM CAPS pH 10.4 into a vessel appropriate for the dispenser used.

b) Retrieve enough 0.5 mg tubes of BODIPY starch to produce the required volume of 2X substrate buffer [calculated in step a) above] at a final concentration of 20-30 ug/mL.

c) Dissolve each 0.5 mg tube in 50 uL DMSO at room temperature, protected from light, with frequent vortexing. This takes more than 15 minutes; some production lots of BODIPY starch dissolve better than others.

d) Add 50 uL 100mM CAPS buffer pH 10.4 to each tube and mix by vortexing.

e) Pool the contents of all tubes and remove any undissolved aggregates by centrifuging for 1 minute at maximum speed in a microfuge.

f) Add the supernatant to the rest of the 100 mM CAPS buffer measured in step a) above.

g) Protect the 2X substrate buffer from light by wrapping in foil.

4. Take plates and substrate buffer to the automation room and program the robot with the following parameters:

a) dispense 4 uL substrate buffer per well

b) 1st read at 1 hour post-substrate, 2nd read at 9 hours, and third read at 17 hours; with 37°C incubation between reads

c) excitation filter: 485 nm; emission filter: 535 nm

d) set the Spectrafluor gain at 70, or the optimal gain for the batch of 2X substrate buffer prepared.

e) ensure that the incubator used will protect assay plates from light.

Day 3

1. Check pintooled plates for clearings in the bacterial lawn at all positions corresponding to wells on the associated assay plate. Also check for clearings in the red starch in any of the pin positions. If plates containing positive controls were used for pintooling, you should be able to see a large clearing zone in the red background. Be wary of contaminants that also form clearing zones in red starch (see comment "Contaminants That Form Clearing Zones in Red Starch" at end of Example 7).

2. Identify putative hits from the data file produced by the robot computer. The KANAL program produced by Engineering simplifies data analysis. As a rule

of thumb, a putative hit is characterized as a well having signal intensity rising at least 1.5 fold over background.

3. For each putative, remove 2 uL from the well and add to a tube containing 500 uL SM buffer and 50 uL CHCl3. Vortex to mix and store at 4°C. This solution will be referred to hereafter as the 4e-3 stock. The original screening plates should be stored at 4°C, protected from light, at least until breakouts are complete.

This is the recommended method of breaking out putative hits. It is a liquid phase assay that relies on confirmation of activity on BODIPY starch. Alternatively, putative hits can be plated directly onto solid phase plates containing red starch such that 2,000-3,000 pfu per hit are examined for clearing zones. However, inability to observe clearing zones on red starch is not necessarily an indication that a putative hit was a false positive. It would then need to be assayed using the format in which it was originally identified (i.e., liquid phase using BODIPY starch as substrate). In addition, very weak positives are more easily identified using the method detailed below.

Day 1

1. In a sterile 50 mL conical tube, combine 0.5 mL OD5 host cells with 45.5 mL NZY. This will be referred to as the host-medium suspension.
2. For each putative hit to be analyzed, aliquot 1 mL of host-medium suspension into each of 3 three sterile microfuge tubes.
3. Set the 12-channel pipetman in multidispense mode with an aliquot size of 20 uL and an aliquot number of **2x**. Mount the pipetman with a clean set of sterile tips.
4. Pour about 1 mL of host-medium suspension into a new sterile solution basin and load the multichannel pipetman.
5. Dispense 20 uL per well into the last row (row P) of a black 384-well plate (12 channels x 2 = 24 wells). This row will be used later for the controls.
6. Expel the remaining liquid in the tips by touching the tips against the surface of the basin and pressing the RESET button on the pipetman. Lay the pipetman down in a way to prevent contamination of the tips. There is no need to change the tips at this point.
7. Pour the remainder of the fluid in the basin into a waste container (like a beaker) taking care to avoid splash-back contamination.

8. For the first putative to be analyzed, take 111 uL of the 4e-3 stock (see Day 2 in *Lambda Microtiter Screening for Amylases*) and add it to the first in a set of three tubes containing 1 mL host-medium suspension (step 2). Vortex to mix. This is *Dilution A*.
9. Take 111 uL of Dilution A and add to the next tube in the set. Vortex to mix. This is *Dilution B*.
10. Take 111 uL of Dilution B and add to the last tube in the set. Vortex to mix. This is *Dilution C*. You should now have three dilutions of phage, where concentrations of each differ by a factor of 10.
11. Pour the contents of Dilution C (the most dilute of the 3 samples) into the solution basin and load the multichannel pipetman.
12. Dispense 20 uL per well into the first row of the 384-well plate (12 channels x 2 = 24 wells).
13. Expel the remaining liquid in the tips by touching the tips against the surface of the basin and pressing the RESET button on the pipetman. Lay the pipetman down in a way to prevent contamination of the tips. There is no need to change the tips at this point.
14. Empty the basin as described above.
15. Pour the contents of Dilution B into the same basin and load the multichannel pipetman.
16. Dispense 20 uL per well into the second row of the 384-well plate.
17. Perform steps 13-16 similarly to dispense Dilution A into the third row of the plate.
18. After all three dilutions have been arrayed into the first 3 rows of the plate, discard all tips and the solution basin into the biohazardous waste container.
19. Mount the pipetman with a clean set of sterile tips and open a new sterile solution basin.
20. Repeat steps 8-19 for each remaining putative hit, using remaining rows on the plate up to row O. Five putative hits can be analyzed on one 384-well plate, with the last row (row P) saved for the controls.
21. Add 0.5 uL of each control to a separate well. Use at least 2-3 separate controls, preferably covering a range of activity.
22. Incubate assay plates at 37°C overnight in a humidified (≥95%) incubator.

Day 2

1. Pintool all breakout plates onto a host lawn with red starch using the same method described for Day 2 in *Lambda Microtiter Screening for Amylases*, except that a 384 position pintool is used.
2. Prepare the 2X BODIPY starch substrate buffer as follows:
   a) Calculate the total volume of 2X substrate buffer solution needed for all breakout plates at 20 uL per well (including any extra deadspace volume required by the dispenser) and measure this amount of 100 mM CAPS pH 10.4 into a vessel appropriate for the dispenser used.
   b) Retrieve enough 0.5 mg tubes of BODIPY starch to produce the required volume of 2X substrate buffer [calculated in step a) above] at a final concentration of 20-30 ug/mL.
   c) Dissolve each 0.5 mg tube in 50 uL DMSO at room temperature, protected from light, with frequent vortexing. This takes more than 15 minutes; some production lots of BODIPY starch dissolve better than others.
   d) Add 50 uL 100mM CAPS buffer pH 10.4 to each tube and mix by vortexing.
   e) Pool the contents of all tubes and remove any undissolved aggregates by centrifuging for 1 minute at maximum speed in a microfuge.
   f) Add the supernatant to the rest of the 100 mM CAPS buffer measured in step a) above.
   g) Protect the 2X substrate buffer from light by wrapping in foil.
3. Dispense 20 uL per well into all breakout plates.
4. Wrap all plates in aluminum foil and incubate at room temperature for 2-6 hours.
5. Read each plate in the Spectrafluor with the following settings:
   a) fluorescence read (excitation filter: 485 nm; emission filter: 535 nm)
   b) plate definition: 384 well black
   c) read from the top
   d) optimal gain
   e) number of flashes: 3
6. On the resulting Excel spreadsheet, chart each putative's 3 rows in a separate graph and check for activity. Ensure that the positives controls produced signals over background.

7. For each putative that appears to have a real signal among the wells, harvest a sample from a positive well as follows:
   a) Select a positive well from a row representing the highest initial dilution.
   b) Transfer 2 uL from that well into a tube containing 500 uL SM and 50 uL $CHCl_3$. This is referred to as the breakout stock.
   c) Store at 4°C.
8. Using methods previously described, plate about 10 uL of each breakout stock onto 150 mm NZY plates using red starch. The objective is to obtain several (at least 20) well-separated plaques from which to core isolates.

Day 3

1. Check pintooled plates for an acceptable incidence of clearings in the bacterial lawn corresponding to wells on the associated assay plate. Also check for clearings in the red starch in the positive controls and in any tested putatives. Be wary of contaminants that also form clearing zones in red starch (see below).
2. From the solid phase plates containing dilutions of breakout stocks, core several isolated plaques, each into 500 uL SM with 50 uL $CHCl_3$. This is referred to as the isolate stock.
3. The isolate stocks can then be individually tested on BODIPY starch using methods described above. This step can be skipped if the plaque that was cored in step 2 produced a clearing zone in the red starch background. The isolate stocks were then be individually tested on BODIPY starch using methods described above. However, this step may be skipped if the plaque that was cored in step 2 produced a clearing zone in the red starch background.

Excisions

Day 1

1. In a Falcon 2059 tube, mix 200 uL OD1 XL1-Blue MRF' host, 100 uL lambda isolate stock and 1 uL ExAssist phage stock.
2. Incubate in 37°C shaker for 15 minutes.
3. Add 3 mL NZY medium.
4. Incubate in 30°C shaker overnight.

Day 2

1. Heat to excision tube to 70°C for 20 minutes.

2. Centrifuge 1000 x g for 10 minutes.
3. In a Falcon 2059 tube, combine 50 uL supernatant with 200 uL EXP505 OD1 host.
4. Incubate in 37°C shaker for 15 minutes.
5. Add 300 uL SOB medium.
6. Incubate in 37C shaker for 30-45 minutes.
7. Plate 50 uL on large LB_{Kan50} plate using sterile glass beads. If the plates are "dry", extra SOB medium can be added to help disburse the cells.
8. Incubate plate at 30°C for at least 24 hours.
9. Culture an isolate for sequencing and/or RFLP.

Growth at 30°C reduces plasmid copy number and is used to mitigate the apparent toxicity of some amylase clones.

Contaminants That Form Clearing Zones in Red Starch

When using red starch on solid medium to assay phage for amylase activity, it is common to see contaminating colony forming units (cfu) that form clearing zones in the red starch. For pintooled plates, it is important to distinguish amylase-positive phage clones from these contaminants whenever they align with a particular well position. The source of the contaminating microbes is presumably the 2% red starch stock solution, which cannot be sterilized by autoclaving or by filtering after preparation. It is thought that they are opportunistic organisms that survive by metabolizing the red starch. In order to reduce these contaminants, use sterile technique when making 2% red starch solutions and store the stocks either at 4°C or on ice.

EXAMPLE 8: Bioinformatic Analysis

The following example describes an exemplary method for determining if a polypeptide is within the scope of the invention, e.g., by Bioinformatic Analysis.

An initial bioinformatic analysis was made with the known hyper-thermophilic α-amylase sequences. Figure 14a shows an alignment of the sequences some of which have been deposited at the NCBI database. This analysis revealed the potential for designing degenerate primers to PCR the entire gene minus its signal sequence (see Figure 14a), yielding potentially novel full-length alpha amylases from a library.

The following libraries were screened by PCR from genomic DNA:

Table 6:

Library #	Name	PCR positive	Subclon ed
5	*A. lithotropicus*	No	
13	*Pyrodictium occultum*	No	
17	*Pyrodictium TAG11*	No	Yes
113	*Deep sea enrichment*	Yes	Yes
170	*Deep sea enrichment*	Yes	Yes
198	*Archaeglobus*	No	
206	*Acidianus sp*	No	
453	*Mixed iceland enrich*	No	
455	*Mixed iceland enrich*	Yes	Yes

Figure 14b shows an alignment of the identified sequences and Table 7, illustrated in Figure 18, lists their relative percent identities.

The amino acid identity ranges from about 85-98% identity. Accordingly, these sequences are useful in shuffling of genes as described herein.

Figure 14c shows the nucleic acid alignment of the corresponding polypeptide sequences above. Expression of these amylases in the expression vector pSE420 and the host cell line XL1-Blue showed 1703 and 1706 to have amylase activity.

EXAMPLE 9: Characterization of Library 63 GP-1 alpha amylase pH optimum and specific activity determination

The following example describes an exemplary method for determining if a polypeptide is within the scope of the invention, e.g., by alpha amylase activity pH optimum and specific activity determination.

In initial experiments, the SEQ ID NO: 81 from *Thermococcus* showed that it was effective in both starch liquefaction for corn wet milling and desizing for textiles. This enzyme has a pH optimum of 4.5 to 5.0. At this lower pH, it is possible to use little or no calcium which lowers overall operating costs and less byproduct formation. In addition, at this low pH, there is decreased chemical usage and ion

exchange load. The industry standard B. licheniformis amylase is suboptimal in both thermostability and pH optimum. The 63GP-1 amylase has a higher application specific activity compared to B. licheniformis amylase and therefore much less enzyme is required to hydrolyze a ton of starch (as much as 20-fold less enzyme can be used).

The pH optimum for the hydrolysis of starch was determined by reacting 50 uL of the GP-1, 0.35 U/ml, with a 100ml of 1% soluble starch solution (0.0175U/g of starch) for 30 minutes at 95 degrees C. The reducing ends generated in the liquefied starch solution were measured by the neocupronine assay, described herein. The percent hydrolysis of cornstarch was determined by measuring the number of sugar reducing ends produced with the neocupronine assay. Seventy grams of buffer solution (pH4-7) was weighed and 100ppm of calcium was added. Thirty grams of cornstarch was mixed into the buffer solution to form a starch slurry. The enzyme was added and the vessels sealed and incubated at 95 degrees C for 30 minutes with an initial heating rate of six degrees C per minute. A 1 ml sample was extracted from the reaction beakers and analyzed by the neocupronine assay. The optimum for GP-1 was between pH 4.5 and 5 , while the commercial *B. licheniformis* amylase performed optimally at about pH 6.0.

EXAMPLE 10: Amylase Ligation Reassembly

The following example describes, inter alia, exemplary methods for determining if a polypeptide is within the scope of the invention, e.g., by the assays described below.

Assay Using RBB-starch

75µl of RBB-starch substrate (1% RBB-insoluble corn starch in 50mM NaAc buffer, pH=4.5) was added into each well of a new 96-well plate (V-bottom). Five micro-liters of enzyme lysate was transferred into each well with substrate using Biomek or Zymark. The plates were sealed with aluminum sealing tape and shaken briefly on the shaker. The plates were incubated at 90^{0}C for 30 minutes, followed by cooling at room temperature for about 5 to 10 minutes. One hundred micro-liters of 100% ethanol was added to each well, the plates sealed and shaken briefly on the shaker. The plates were then centrifuged 4000 rpm for 20 minutes using bench-top centrifuge. 100µl of the supernatant was transferred into a new 96-well plate (flat bottom) by Biomek and read OD_{595}. Controls: SEQ ID NO:81, SEQ ID NO:77, SEQ ID NO:79.

Assay using FITC-starch

Added 50µl of substrate (0.01% FITC-starch in 100mM NaAc buffer, pH=4.5) into each well of a new 384-well plate. Transferred 5µl of enzyme lysate into each well with substrate and incubated the plate at room temperature overnight. The polarization change of the substrate, excitation 485nm, emission 535nm, was read for each well. Controls: SEQ ID NO.: 81, SEQ ID NO.: 77, SEQ ID NO.: 79. Preferably 96 well plates are used for all assays.

Confirmation of new active clones

Each positive clone from screening was grown and induced using a standard protocol. Each clone was examined for growth (i.e., cell density over time), activity at per cell level (RBB-starch assay and liquefaction assay), expression (protein gel) and solubility of protein (by microscope analysis). The confirmed new elevated clones were transferred for fermentation.

Example 11: Exemplary protocol for liquefying starch and measuring results

The following example described and exemplary protocol for liquefying starch using selected amylases of the invention.

Amylases having a sequence as set forth in SEQ ID NO:10 and SEQ ID NO:4 demonstrated activity on liquefied starch at pH 4.5 or 6.5 using the reaction conditions show below.

Reaction Conditions: 100 mM PO_4 pH 6.5, 1% (w/w) liquefied starch DE 12 at 55°C. Both TLC and HPLC assays were done to verify activity. The data from both assays showed that the clones were active.

pH profiles for the amylases having a sequence as set forth in SEQ ID NO:4 and SEQ ID NO:10 were run using phosphate buffer pHed from 3.0 - 6.5, at 55°C. From the amount of observable hydrolysis, it could be visually said that the clones were more active at certain pH values than at other values at the above indicated reaction conditions:

SEQ ID NO:4 - active from pH 5.0 - 6.5

SEQ ID NO:10 - active from pH 4.5 - 6.5

An exemplary protocol for the saccharification of liquefied starch at pH 6.5:

- Adjust the pH of the liquefied starch to the pH at which the saccharification(s) will be performed. Liquefy starch in 100 mM sodium acetate buffer, pH 4.5 with 100 mM sodium phosphate salts added so that before saccharification, the pH could be adjusted to pH 6.5.
- Weigh 5 gram samples of liquefied starch into tared bottles.
- Use 0.04% (w/w) Optidex L-400 or approximately 400 mL of 1-10 diluted stock Optidex L-400 per 100 grams of liquefied starch.
- Calculate the milligrams of Optidex L-400 contained in the 400 mL of 1-10 diluted stock Optidex L-400. Next, calculate the volume of lysates needed to give the same concentration of enzyme as the Optidex L-400.
- Add enzymes to liquefied starch samples and incubate at desired temperature (50C°). After 18 hours determine DE and prepare a sample for HPLC analysis.

An exemplary DE Determination:

Exemplary Neocuproine Assay:

A 100ml sample was added to 2.0ml of neocuproine solution A (40g/L sodium carbonate, 16g/L glycine, 0.45g/L copper sulfate). To this was added 2.0 ml of neocuproine solution B (1.2g/L neocuproine hydrochloride-Sigma N-1626). The tubes were mixed and heated in a boiling water bath for 12 minutes; cooled, diluted to 10ml volume with DI water and the OD read at 450nm on the spectrophotometer. The glucose equivalent in the sample was extrapolated from the response of a 0.2mg/ml glucose standard run simultaneously.

Exemplary HPLC Analysis:

Saccharification carbohydrate profiles are measured by HPLC (Bio-Rad Aminex HPX-87A column in silver form, 80°C) using refractive index detection. Mobile phase is filtered Millipore water used at a flow rate of 0.7 ml/min. Saccharification samples are diluted 1-10 with acidified DI water (5 drops of 6 M HCl into 200 mL DI water) then filtered through a 0.45 mm syringe filter. Injection volume is 20 uL.

Exemplary TLC:

Reaction products were w/d at hourly timepoints and spotted and dried on a TLC plate. The Plate was then developed in 10:90 water:isopropanol and visualized

with either a vanillin stain or CAM stain and then heated to show reducible sugars. The liquefied starch was partially hydrolyzed to glucose in cases where activity was observed.

EXAMPLE 12: Starch Liquefaction using Amylases of the Invention

This example describes an exemplary method of the invention for liquefying starch using amylases of the invention.

Amylase concentrate was prepared from fermentation broths by heat treatment, cell washing, alkaline extraction using microfiltration and ultrafiltration (48% overall yield). The UF concentrate was neutralized with acetic acid and formulated with 30% glycerol at pH 4.5. The activity level of the slurry formulation was representative of a commercial product ($120U^1/g$ – 0.5kg/ ton starch).

Standard Amylase Activity Assay

A 1 mL cuvette containing 950 µL of 50 mM MOPS pH 7.0 containing 5 mM PNP-α- D—hexa-(1→4)-glucopyranoside was placed in the Peltier temperature controller of the Beckman DU-7400 spectrophotometer preheated to 80°C. The spectrophotometer was blanked at 405nm and 50 µL of the enzyme solution was added to the cuvette, mixed well and the increase in the OD_{405nm} was monitored over a one-minute interval. The $\Delta OD_{405nm/min}$ rate is converted to a standard unit of µmole/minute from the OD_{405nm} response of 50 µL of 1 µmole/mL PNP in 950 mL 50 mM MOPS at pH 7.0 - 80°C. One standard Diversa unit of thermostable alpha amylase (DTAA) is equal to the amount of enzyme that will catalyze the release of 1 µmole/mL/minute of pNP under the defined conditions of the assay.

Standard Glucoamylase Activity Assay

A 1 mL cuvette containing 950 µL of 50 mM MOPS pH 7.0 containing 5 mM pNP-α- D-glucopyranoside was placed in the Peltier temperature controller of the Beckman DU-7400 spectrophotometer preheated to 60°C. The spectrophotometer was blanked at 405nm and 50 µL of the enzyme solution was added to the cuvette, mixed well and the increase in the OD_{405nm} was monitored over a one-minute interval. The ΔOD_{405nm}/min rate is converted to a standard unit of µmole/minute from the OD_{405nm}

response of 50 µL of 1 µmole/mL pNP in 950 mL 50 mM MOPS at pH 7.0-60°C. One standard Diversa unit of glucoamylase (DGA) is equal to the amount of enzyme that will catalyze the release of 1 µmole/mL/minute of pNP under the defined conditions of the assay.

Dextrose Equivalent Determination

The neocuproine method was used to measure the DE. Selected samples were measured by both the Invention procedure and by a GPC analyst using the GPC Fehlings procedure.

Neocuproine Assay

A 100 µl sample was added to 2.0 ml of neocuproine solution A (40 g/L sodium carbonate, 16g/L glycine, 0.45g/L copper sulfate). To this was added 2.0 ml of neocuproine solution B (1.2 g/L neocuproine hydrochloride-Sigma N-1626). The tubes were mixed and heated in a boiling water bath for 12 minutes; cooled, diluted to 10ml volume with DI water and the OD read at 450 nm on the spectrophotometer. The glucose equivalent in the sample was extrapolated from the response of a 0.2mg/ml glucose standard run simultaneously.

The starch sample is diluted ~1 to 16 with DI water with the exact dilution recorded. Ten milliliters of the diluted sample was added to 20 mls of DI water. Ten milliliters of Fehlings solution A and B were added to the diluted starch. The sample was boiled for 3 minutes and cooled on ice. Ten milliliters of 30% KI and 10ml of 6N H_2SO_4 was added. The solution was titrated against 0.1N sodium thiosulfate. The titrant volume is recorded and used to calculate the DE.

Residual Starch Determination

Post-saccharification samples were checked for residual starch using the Staley iodine procedure.

Twenty grams of sample was weighed into a large weigh dish. 45 µL of Iodine solution is added to the weigh dish and the starch solution is mixed well. Dark blue indicates the presence of starch, a light blue-green indicates slight starch, light green indicates a trace of starch and yellow-red, absence of starch. Iodine solution is prepared by dissolving 21.25 grams of iodine and 40.0 grams of potassium iodide in one liter of water.

Oligosaccharide Profile

Liquefaction and saccharification carbohydrate profiles were measured by HPLC (Bio-Rad Aminex HPX-87C column in calcium form – 80°C) using refractive index detection.

Gel Permeation Chromatography

The molecular weight distribution was determined by chromatography on a PL Aquagel-OH column with mass detection by refractive index (Waters Model 2410). A Viscotek Model T60 detector was used for continuous viscosity and light scattering measurements.

Capillary Electrophoresis

Beckman Coulter P/ACE MDQ Glycoprotein System – separation of APTS derivatized oligosaccharides on a fused silica capillary - detection by laser-induced fluorescence.

Primary Liquefaction

Line starch directly from the GPC process is pumped into a 60 liter feed tank where pH, DS (dry solids) and calcium level can be adjusted before liquefaction. The amylase is added to the slurry. The 32% DS slurry is pumped at 0.7 liter/minute by a positive displacement pump to the jet - a pressurized mixing chamber where the starch slurry is instantaneously heated to greater than 100C by steam injection. The gelatinized partially liquefied starch is pumped through a network of piping (still under pressure) to give the desired dwell time (5 minutes) at temperature. The pressure is released into a flash tank and samples can be taken. Samples were taken in duplicate.

Secondary Liquefaction

The liquefied starch was collected in one liter glass bottles and held in a water bath at 95C for 90 minutes.

Saccharification

Liquefied starch was cooled to 60C, the pH adjusted to 4.5 and the samples treated with glucoamylase. Saccharification progress was monitored over time by HPLC.

Saccharification

The liquefied syrups produced with each amylase were adjusted to approximately pH 2.5 with 6N HCl immediately after the 90 minute secondary liquefaction to inactivate any residual amylase. The syrups were then adjusted to pH 4.5, placed in a 60°C water bath and saccharified with three levels of glucoamylase. The extent of saccharification was monitored by HPLC at 18-88 hour time points.

The liquefied syrups were saccharified with the standard dosage – 0.04% of a double-strength glucoamylase - and two lower dosages (50% and 25%) to monitor any differences in the saccharification progress.

Saccharification Progress - % dextrose development vs time – 0.04% glucoamylase

Amylase	18 hr	24 hr	40 hr	44 hr	88 hr
Commercial	70.2	78.4	86.1	86.7	94.2
SEQ ID NO:437	79	88.6	92.5	92.8	95.3
SEQ ID NO:6	74.1	85.9	91.9	91.6	94.8

Saccharification Progress - % dextrose development vs time – 0.02% glucoamylase

Amylase	18 hr	24 hr	40 hr	44 hr	88 hr
B.licheniformis Amylase	54.5	66.7	76.1	77.2	90.9
SEQ ID NO:437	60.1	72	84.8	85.3	93.6
SEQ ID NO:6	57.1	70	84	86.5	92.5

Post-Saccharification sugar profile

In these studies and all previous saccharification studies, the final glucose level achieved after saccharification by amylases of the invention and *B. licheniformis* in liquefied syrups is essentially identical. The DP2 (maltose) level is also similar. These large fragments are poor substrates for glucoamylase and tend to be converted slowly, if at all, into smaller fragments and ultimately, glucose.

	Glucose	DP2	DP3	>DP7
SEQ ID NO:437	95.25	2.39	1.13	0.91
Commercial	94.16	2.10	0.39	2.91
SEQ ID NO:6	94.77	2.27	1.48	0.82

Molecular weight distribution

The molecular weight distribution of syrups liquefied to DE's of 12 and 18 by the exemplary amylases of the invention SEQ ID NO:6 and SEQ ID NO:437, and commercial *Bacillus licheniformis* and commercial *Bacillus stearothermophilus*, were measured by gel permeation chromatography using detection by refractive index, light scattering and viscosity. Both the *B. licheniformis* and *B. stearothermophilus* amylases generate a bimodal distribution – the primary peak centered at 2000, a secondary peak at 32,000 with a shoulder extending past the 160,000 range. The lower molecular weight peak represents approximately 60% of the total mass of the sample. The exemplary amylases of the invention exhibit a single peak at 2000 with very little above 30,000.

HPLC

The DE 12 and 18 syrups produced by the exemplary amylases of the invention SEQ ID NO:6 and SEQ ID NO:437 and commercial *Bacillus licheniformis* and commercial *Bacillus stearothermophilus* amylases were analyzed by HPLC. Both techniques produce fingerprints characteristic of each class of amylase; the oligosaccharide patterns are different for *B. licheniformis* amylase vs *B. stearothermophilus* amylase vs the exemplary amylases of the invention. The liquefied syrups of the invention (e.g., syrups made by methods of the invention and/or made by enzymes of the invention) exhibit evidence of greater branching in the oligosaccharides.

HPLC only resolve the oligosaccharides in the <DP15 range – larger fragments are not visible in these techniques. Bacillus amylases are known to liquefy starch in a manner such that the amylopectin fraction is hydrolyzed less extensively than the amylose fraction. These >DP30 amylopectin fragments are contained in the high molecular weight fraction centered at 32,000 and consequently, little evidence of branching is seen in the HPLC analyses of the Bacillus liquefied syrups. The <DP15 oligosaccharides from Invention amylases contain fragments from both amylose and amylopectin.

EXAMPLE 13: Starch Liquefaction at acidic conditions using amylases of the invention

The invention provides methods for liquefying starch using amylases of the invention, including amylases active under acidic conditions, e.g., between about pH 4.0 and 5.0, e.g., pH 4.5. The conversion of starch to glucose can be catalyzed by the sequence action of two enzymes: alpha-amylases of the invention to liquefy the starch (e.g., the hydrolysis of high molecular weight glucose polymers to oligosaccharides consisting of 2 to 20 glycose units, typically a dextrose equivalent of 10 to 12, by an amylase of the invention), followed by saccharification with a glycoamylase (which can be a glycoamylase of the invention). In one aspect, processing is in a corn wet milling plant producing a starch slurry having a pH or about 4.0 to 4.5. In one aspect, the pH is raised, e.g., to 5.8 to 6.0 before liquefaction to accommodate an alpha amylase with a low pH activity and stability (which can be an alpha amylase of the invention). In one aspect, amylases of the invention can liquefy starch at pH 4.5 to dextrose equivalents ranging from 12 to 18; in one aspect, using alpha amylases of the invention at levels of about 3 to 6 grams per ton of starch. In this aspect, use of alpha amylases of the invention enables starch liquefaction to be conducted at pH 4.5.

In one aspect, starch liquefaction is conducted at pH 4.5 for 5 minutes at 105°C to 90 minutes at 95°C using amylases of the invention. The quantity of enzyme was adjusted in order to adjust a target DE of 12 to 15 after liquefaction. In one aspect, the liquefied starch is then saccharified with a glucoamylase, e.g., an *Aspergillis* glucoamylase, for about 48 hours at about pH 4.5 and 60°C. If the saccharified syrup did not contain at least 95% glucose, the target liquefaction DE was raised and the saccharification repeated until the liquefaction eventually did produce a saccharified syrup containing more than 95% glucose. The amylase protein required to produce a suitable liquefied feedstock for saccharification was determined by PAGE.

EXAMPLE 14: Starch Liquefaction using amylases of the Invention

This example describes an exemplary method for liquefying starch using amylases of the invention and characterizes liquefaction oligosaccharide patterns of the exemplary enzymes of the invention SEQ ID NO:6 and SEQ ID NO:437 (encoded by SEQ ID NO:436) vs commercial *Bacillus licheniformis* and *Bacillus stearothermophilus* amylases. These results compare the saccharification progress and final dextrose levels from syrups generated by enzymes of the invention and commercial amylases.

Three commercial enzymes, Genencor Spezyme AA, and two others all required more than double the recommended dosage to achieve the target Dextrose equivalent (DE). Dextrose equivalent (DE) is the industry standard for measuring the concentration of total reducing sugars, calculated as D-glucose on a dry weight basis. Unhydrolyzed granular starch has a DE of virtually zero, whereas the DE of D-glucose is defined as 100.

These results confirm the "double dosage" effect for all *Bacillus* amylases and gives more credence to the proposal that the observed dosage for SEQ ID NO:437 in the trials is also twice the value which would be required under more normal conditions. The projected "normal" dosage, 60-70 Units/kilo starch at pH 4.5 to reach a 19 DE, is consistent with the laboratory liquefaction data.

The oligosaccharide pattern generated by amylases of the invention is different from the *Bacillus* profiles. The molecular weight distribution for the *Bacillus* amylases (gel permeation chromatography with detection by light scattering and viscosity) is bimodal with a substantial fraction at the very high molecular weight range (>300,000) even at an 18DE. The SEQ ID NO:437 at 18DE exhibits a uniform distribution with nothing greater than 20,000. This is consistent with the lower viscosity for syrups of the invention (e.g., syrups made by methods of the invention, or, made using enzymes of the invention). The DP (degrees of polymerization) profiles as measured by HPLC also reflects this difference in action pattern.

In this study, as well as in the previous studies, the final glucose level after saccharification of amylases of the invention liquefied syrups vs the *Bacillus* syrups is the same for both cases. However, saccharification data from, e.g., GPC studies, confirm that the non-dextrose residuals for the amylases of the invention are different from the *Bacillus* amylase syrups. Although the dextrose and maltose levels are essentially the

same for both, the amylases of the invention have a higher DP3 fraction but lower amount of the "highers" vs. the *Bacillus* enzyme. Consistent with the absence of high molecular weight fragments after liquefaction, the post saccharification syrups of the invention have a lower content of the >DP7 fraction.

	Glucose	DP2	DP3	>DP7
SEQ ID NO:2	95.25	2.39	1.13	0.91
Commercial	94.16	2.10	0.39	2.91
SEQ ID NO:6	94.77	2.27	1.48	0.82

SEQ ID NO:437 amylase concentrate was prepared from fermentation broths by heat treatment, cell washing, alkaline extraction using microfiltration and ultrafiltration (48% overall yield). The UF concentrate was neutralized with acetic acid and formulated with 30% glycerol at pH 4.5. The activity level of the slurry formulation was representative of a commercial product (120U1/g – 0.5kg/ ton starch).

Example 15: Alkaline Amylases for Laundry and Autodishwash Applications

In one aspect, the invention provides detergents comprising amylases of the invention, including amylases active under alkaline conditions, and methods of making and using them.

Three alkali-stable amylase enzymes of the invention were compared to and outperformed a commercial benchmark enzyme with respect to features important in laundry and automatic dishwashing (ADW) applications:

- Amylase having a sequence as set forth in SEQ ID NO:212 (encoded by SEQ ID NO:211) outperformed the purified commercial benchmark enzyme in the ADW wash test on starch-coated slides and was very resistant to hydrogen peroxide.
- Amylase having a sequence as set forth in SEQ ID NO:210 (encoded by SEQ ID NO:209) and SEQ ID NO:212 (encoded by SEQ ID NO:211) outperformed the purified commercial benchmark enzyme in the presence of a laundry/ADW formulation using a soluble substrate.
- In the presence of chelators, amylase having a sequence as set forth in SEQ ID NO:439 (encoded by SEQ ID NO:438) was very stable and

amylase having a sequence as set forth in SEQ ID NO:441 (encoded by SEQ ID NO:440) was moderately stable.

- Amylase having a sequence as set forth in SEQ ID NO:210 (encoded by SEQ ID NO:209) and amylase having a sequence as set forth in SEQ ID NO:212 (encoded by SEQ ID NO:211) and amylase having a sequence as set forth in SEQ ID NO:441 (encoded by SEQ ID NO:440) have very alkaline pH optima in the range of pH 10 to 11. Amylase having a sequence as set forth in SEQ ID NO:445 (encoded by SEQ ID NO:444) and having a sequence as set forth in SEQ ID NO:439 (encoded by SEQ ID NO:438) have pH optima around 8 while retaining significant activity at pH 10.
- Amylase having a sequence as set forth in SEQ ID NO:441 (encoded by SEQ ID NO:440) and having a sequence as set forth in SEQ ID NO:439 (encoded by SEQ ID NO:438) were thermophilic, performing best at 65° to 70°C.

Biochemical characterization

Five amylases of the invention, three with alkaline pH optima, were characterized for pH optimum and temperature optimum, as described in Table 1. "SEQ ID NOS:209, 210" refers to an amylase having a sequence as set forth in SEQ ID NO:110, encoded by SEQ ID NO:209, etc.

Table 1

Amylase	pH optimum	Temp. optimum (°C)*
SEQ ID NOS:209, 210	11	55
SEQ ID NOS:211, 212	10	50
SEQ ID NOS:440, 441	10	70
SEQ ID NOS:444, 445	8	40
SEQ ID NOS:438, 439	8	65

Temperature optima were determined at pH 10 for the amylase having a sequence as set forth in SEQ ID NO:210, encoded by SEQ ID NO:209 ("SEQ ID NOS:209, 210"); SEQ ID NOS:211, 212; and SEQ ID NOS:440, 441 and at pH 8 for SEQ ID NOS:444, 445 and SEQ ID NOS:438, 439.

The pH profiles for amylases of the invention compared to the

benchmark enzyme currently used in a commercial laundry/ADW product are presented in Figure 1. All of the enzymes of the invention demonstrated optimal activity between pH 8 and 10, whereas the commercial benchmark enzyme was most active at pH below 8 and had only residual activity at pH 10. Figure 19 shows the pH profile of the tested amylases of the invention and the commercial benchmark enzyme. Purified protein was added to buffers of the indicated pH containing soluble substrate and the activity was measured. Initial rates were calculated over 10 min and converted to a percentage of the maximum rate.

The temperature profiles of enzymes of the invention are presented in Figure 20. Three were most active between temperatures 45°C and 55°C, while the amylase having a sequence as set forth in SEQ ID NO:441 (encoded by SEQ ID NO:440) ("SEQ ID NOS:440, 441") and SEQ ID NOS:438, 439 had optimum activity between 60°C and 70°C. Figure 20 shows the temperature activity profiles of the tested amylases of the invention. Activity of purified protein was measured at pH 10 (SEQ ID NOS:209, 210, SEQ ID NOS:211, 212, SEQ ID NOS:440, 441) or pH 8 (SEQ ID NOS:444, 445, SEQ ID NOS:438, 439) at the indicated temperature. Activity was measured either by a reducing sugar assay or by monitoring the fluorescence at 520 nm (485 nm excitation) when BODIPY-starch was used. Initial rates were calculated and converted to a percentage of the maximum rate.

Application testing

Experiments were designed to assess the activity and stability of the tested alkaline amylases of the invention in laundry/ADW formulations and with the components individually. Figures 21, 22 and 23 present the results of experiments using a soluble starch substrate. Figure 24 presents results using a solid substrate - the industry-standard starch-coated slides.

Amylase having a sequence as set forth in SEQ ID NO:439 (encoded by SEQ ID NO:438) ("SEQ ID NOS:438, 439") was very resistant to the chelator EDTA (Figure 21) and SEQ ID NOS:211, 212 displayed significant resistance to hydrogen peroxide (Figure 22). In contract, the commercial benchmark enzyme was not functional in the presence of either component under the conditions of the experiments. In the presence of the complete laundry/ADW formulation, SEQ ID NOS:209, 210 and SEQ ID

NOS:211, 212 were much more active on soluble substrate than the commercial benchmark enzyme (Figure 23).

Figure 21 shows enzyme activity in the presence of EDTA. Purified proteins were incubated at 50°C in the presence or absence of 5mM EDTA for the indicated time, after which residual amylase activity was measured using soluble substrate. Activity in the presence of EDTA is expressed as the % of activity in the absence of chelator. Figure 22 shows enzyme activity in the presence of peroxide hydroxide. Purified proteins were incubated at 50°C in the presence or absence of 1M H_2O_2 for the indicated time after which amylase activity was measured using soluble starch. Activity in the presence of peroxide hydroxide is presented as the % of activity in the absence of H_2O_2. Figure 23 shows enzyme activity in the ADW solution (distilled water, hardening solution, bleach, chelators, surfactants) with soluble substrate (BODIPY-starch). Purified proteins reacted with the soluble starch at 40°C in the presence of laundry/ADW formulation. Initial rates were calculated over 5 minutes and expressed as fluorescent units (FU)/s per ng of protein.

The lead performers emerging from the tests on soluble substrate were the amylase having a sequence as set forth in SEQ ID NO:210 (encoded by SEQ ID NO:209) ("SEQ ID NOS:209, 210") and SEQ ID NOS:211, 212. These amylases, along with SEQ ID NOS:440, 441, were compared with the commercial benchmark enzyme in the industry-standard wash test on the starch-coated slides. Results of these experiments are presented in Figure 24. The enzyme having a sequence as set forth in SEQ ID NO:212 (encoded by SEQ ID NO:211) consistently outperformed the purified benchmark enzyme in this test although the formulated benchmark enzyme showed better performance. The nature of the benchmark commercial formulation is unknown, but the purified benchmark enzyme displayed two-fold increase in activity in the presence of Bovine Serum Albumin (BSA). Figure 24 shows the results of the wash tests with starch-coated slides. Purified proteins were incubated with slides at 50°C for 30 min in the presence of ADW solution (distilled water, water hardening solution, bleach, chelators, surfactants). Starch removal was measured comparing weight loss after the enzyme treatment to the initial weight of the slide.

Summary of the characterization of exemplary amylases

The gene encoding the amylase having a sequence as set forth in SEQ ID

NO:212 (encoded by SEQ ID NO:211) ("SEQ ID NOS:211, 212") was isolated from an environmental library collected from a biotope with a pH of 11.0 and temp of 41°C. The amylase encoded by this gene belongs to Family I and does not contain any known Starch/Carbohydrate Binding Domains. The protein has been expressed with and without a C-terminal histidine tag, and in non-glycosylating and a glycosylating host. Enzyme expressed in all of these Host/His tag combinations have pH optima around 10 and temperature optima around 50°C (experiments represented by Figures 19 and 20). The enzyme expressed in the glycosylating host with a His tag was used for the experiments represented by Figures 21 through 24. The presence of the His tag does not seem to affect specific activity, however, glycosylation appears to result in a slightly lower specific activity than that without glycosylation.

In summary:

- The best performer in these application assays was the amylase having a sequence as set forth in SEQ ID NO:212 (encoded by SEQ ID NO:211) ("SEQ ID NOS:211, 212").
- pH and temperature optima of SEQ ID NOS:211, 212 meet the requirements for laundry/ADW applications and SEQ ID NOS:211, 212, with proper formulation, should exceed the performance of the commercial benchmark enzyme.

Example 16: Identification and characterization of a thermostable glucoamylase

The following example describes the identification and characterization of an exemplary thermostable amylase of the invention having glucoamylase activity.

Nucleic Acid Extraction: The filamentous fungus *Thermomyces lanuginosus* ATCC 200065 was grown in liquid culture in Potato Dextrose Medium (Difco, BD, Franklin Lakes, NJ). Biomass was collected and high molecular weight genomic DNA was isolated using DNEASY™ (DNeasy) Plant Maxi Kit (Qiagen, Valencia, CA) using standard protocols. Total RNA was also isolated using RNEASY™ (RNeasy) Plant Mini Kit (Qiagen) using standard protocols.

Library Construction: *Thermomyces* genomic DNA was partially digested with restriction enzymes and fragments between 1-10 kb were purified for construction of a genome library. The fragments were ligated into the vector Lambda Zap Express™ (Stratagene, San Diego, CA) and packaged into infectable phage as per manufacturer's instructions.

Library Screening: The above lambda library was used to infect XL1 Blue MRF' cells (Stratagene) in top agar. Approximately 50,000 pfu of phage was added to 600 ul of cells OD600=1. The mixture was incubated at 37°C for 15 minutes in a water bath and then added to 6 ml melted 0.7% top agar and plated onto NZY agar plates. The plate was then incubated overnight at 39°C. A nylon circle (F. Hoffmann-La Roche Ltd., Basel Switzerland) was laid on top of the resulting plaque lawn and lifted back up with some of the phage adhering to the nylon. The nylon was submerged in 1.5M NaCl, 0.5M NaOH for 2 minutes, 1.5M NaCl, 0.5M Tris pH 7.6 for 5 minutes and 2X SSC, 0.2M Tris pH7.6 for 30 seconds. The nylon filter was then UV crosslinked in a Stratagene crosslinker.

A 639 bp PCR fragment from the glucoamylase gene of *Aspergillus niger* was generated from *Aspergillus* genomic DNA for use as a probe. The primers (5'-GCGACCTTGGATTCATGGTTGAGCAAC-3' (SEQ ID NO:595) and 5'-CACAATAGAGACGAAGCCATCGGCGAA-3') (SEQ ID NO:596) were used in the PCR reaction that utilized the Expand High Fidelity PCR Kit™ (Roche) using 30 cycles of 95°C for 20 seconds, 55°C for 30 seconds, and 72°C for 1 minute in a thermal cycler. This PCR fragment is composed of exons 1-4 of the *Aspergillus* glucoamylase gene. The isolated PCR fragment was prepared as a radioactive probe using the Prime It Kit™ (Stratagene) following manufacturer's instructions.

The library filter lifts were washed in a prehybridization solution (DIG Easy Hyb™, Roche) for two hours at 42°C in a hybridization oven (Robbins). The probe was added to 15ml fresh DIG Easy Hyb™ and used to replace the prehybridization solution. The filter was washed with probe overnight at 45°C. The probe was then removed and the filter washed once with 2X SSC, 0.1% SDS for 15 minutes, and twice with 0.1X SSC, 0.1% SDS for 15 minutes each. The nylon filter was then exposed to x-ray film overnight at –80C. Following developing, hybridization spots on the x-ray film were used to identify clones from the original plate. An agar plug was taken from the plate where the spots lined up and suspended in SM buffer to release the phage into solution. Several isolated plaques corresponding to *Thermomyces* genomic fragments containing all or part of the glucoamylase gene were thus isolated.

100 ul of isolated phage stock was added to 200ul XL-1 Blue MRF' cells (Stratagene) and 1 ul ExAssist™ helper phage (Stratagene). The mixture was incubated at 37C for 15 minutes, and 3 ml of 2X YT media was added. This was then incubated at

37°C with shaking for 2.5 hours. The mix was then heated for 20 minutes at 70°C and cooled on ice. 100 ul of the mix was removed and added to 200 ul SOLR cells (Stratagene) and incubated at 37C for 15 minutes. 50 ul was plated on LB kanamycin (50 ug/ml) plates and incubated overnight at 37°C. Resulting colonies contain cloned genomic fragments in the plasmid pBK-CMV.

Sequencing: DNA sequencing on candidate clones were performed with the BigDye Terminator Cycle Sequencing Version 2.0 Kit™ (Applied Biosystems, Foster City, CA) and a 3700 DNA Analyzer™ (Applied Biosystems) using manufacturer's protocols. A genomic clone was identified with a 4.1 kb insert that contained the entire glucoamylase gene and flanking sequence, as set forth in SEQ ID NO:587. Potential introns were identified by comparing this sequence with consensus sequences for introns in *Aspergillus*. The *Thermomyces lanuginosus* nucleotide sequence has an open reading frame encoding a protein of 617 amino acids, interrupted by four introns of 64 bp, 61 bp, 80 bp, and 57 bp respectively.

cDNA Synthesis: The primers 5'-ATGTTATTCCAACCGACTTTGTGCGC-3' (SEQ ID NO:597) and 5'-TCATCGCCACCAAGAATTCACGGTG-3' (SEQ ID NO:598) were used in a cDNA synthesis reaction using a Thermoscript rtPCR Kit™ (Invitrogen) using manufacturer's protocols. The template for synthesis was total RNA isolated from *Thermomyces lanuginosus* cells growing on potato dextrose media (Difco). An 1854 bp fragment from the reaction was isolated, cloned and sequenced, with the nucleic acid sequence set forth in SEQ ID NO:593.

Expression Cloning: Primers were designed for overexpression of Thermomyces glucoamylase in the host Pichia pastoris. The primers 5'-GTCTCGAGAAAAGAGCAACGGGCTCGCTCGAC-3' (SEQ ID NO:599) and 5'-GTTCTAGATCATCGCCACCAAGAATTCACGGT-3' (SEQ ID NO:600) were used to generate a PCR fragment using the cDNA clone as a template using 30 cycles of 95°C for 20 seconds, 55°C for 30 seconds, 72°C for 2 minutes, using Expand High Fidelity PCR Kit™ (Roche) and manufacturer's protocols. The PCR fragment was digested with the restriction enzymes Xho I and Xba I and ligated into the corresponding restriction sites of the plasmid pPIC Z alpha (Invitrogen). The construct was transformed into *Pichia pastoris* Strain X-33 (Invitrogen) where the construct integrates stably into the *Pichia* chromosome. Selection was based on resistance to zeocin. This construct was designed

such that the *Pichia* clone can be induced with methanol to secrete the mature *Thermomyces* glucoamylase into the media.

A 1-liter culture of the *Pichia* expression clone was inoculated with an overnight starter culture in BMGY and grown overnight at 30°C in a shake flask. The cells were collected by centrifugation the following day and resuspended in 1 liter of BMMY. The cells were cultured at 30°C in a shake flask for 3 days with methanol added to 0.5% final every 24 hours. The media containing the expressed glucoamylase enzyme was then collected and tested in a glucoamylase activity assay and SDS PAGE electrophoresed using standard protocols to determine the protein size.

Primers were also designed for overexpression of the *Thermomyces* glucoamylase gene in *Escherichia coli*. The primers (SEQ ID NO:601) 5'-GTCCATGGCAACGGGCTCGCTCGAC-3' and (SEQ ID NO:602) 5'-GTTCTAGATCATCGCCACCAAGAATTCACGGT-3' were used to generate a PCR product as before, from the cDNA template. The PCR fragment was digested with the restriction enzymes Nco I and Xba I and ligated into corresponding restriction sites of the plasmid pSE420 (Invitrogen). The construct was transformed into *Escherichia coli* Strain XL-1 Blue MR (Stratagene). Selection for the plasmid was based on ampicillin resistance. The glucoamylase gene is under the control of the lac-z promoter in this plasmid vector and is induced with IPTG (isopropyl-thio-galactopyranoside). The construct was designed such that the mature glucoamylase gene will be expressed within the *Escherichia* cell and will contain an extra methionine residue at the N-terminus.

Standard assay: Enzyme aliquots were added to a solution of 5 mM buffer, 3 mM malto-oligosaccharides (Sigma, M-3639) in a waterbath. 100 ul aliquots removed at time points to 200 ul glucose oxidase reagent (Sigma, GAGO-20) and incubated 37°C, 30 min. The reaction was stopped with addition of 12 N sulfuric acid and the absorbance at 540 nm determined. The full-length version of the enzyme (SEQ ID NO:594) was tested for pH, temperature and substrate utilization. As noted below, data demonstrated that the pH optimum to be around pH 5.5. Data demonstrated that the enzyme (SEQ ID NO:8) is stable at 70°C with a rapid irreversible loss of activity between 70°C and 75°C. Data demonstrated that the enzyme (SEQ ID NO:594) hydrolyses oligosaccharides down to maltose with the rate of hydrolysis being higher for longer saccharides. The rate in cleaving 1,6 linkages is much slower than 1,4 as observed in the substrate panose which has a 1,6 linkage at the non-reducing end. The catalytic domain version appears to be less

thermostable. The enzyme (SEQ ID NO:594) has a good rate of hydrolysis at 50°C but appears to die at 70°C.

Activity Assay: Enzyme (SEQ ID NO:594) activity was measured by the release of free glucose from an oligo-dextrin substrate. The liberated glucose was then oxidized in a coupled reaction resulting in a colored product. An enzyme (SEQ ID NO:594) aliquot added to solution of 5mM buffer, 3mM malto-oligosaccharides (Sigma, M-3639) in a water bath. 100 ul aliquots removed at time points to 200ul glucose oxidase reagent (Sigma, GAGO-20) and incubated 37°C, 30 min. The reaction was stopped with addition of 12 N sulfuric acid and the absorbance at 540 nm determined. Time points were then plotted to determine the relative rate for the reaction.

pH Profile: Acetate buffer (pH 4.0, 4.5, 5.0, and 5.4) as well as phosphate buffer (pH 6.2, 7.0, 8.1) were used in an activity assay to determine the relative rate for the glucoamylase (SEQ ID NO:594) at each pH. The rates were then plotted, as illustrated in Figure 5. The enzyme (SEQ ID NO:594) appears to have maximal activity around pH 5.5.

Temperature Profile: The relative rate of the enzyme (SEQ ID NO:594) at various temperatures (50°C, 60°C, 70°C, 80°C, and 85°C) was determined in acetate buffer pH 5.3. The rates are plotted in Figure 6. The enzyme (SEQ ID NO:594) appears to have maximal activity at 70°C, above which there is a rapid loss of activity.

Temperature Stability Data: Enzyme (SEQ ID NO:594) was added to 5 mM acetate buffer at the indicated temperature. Enzyme (SEQ ID NO:594) aliquots were removed to ice at 4 minute intervals. The aliquots were then tested for activity on substrate for 20 minutes at 70°C, and the data is illustrated in Figure 7.

Substrate Utilization: The dextrins maltose (G2), maltotriose (G3), panose (Pan), maltotetraose (G4), and maltoheptaose (G7), were substituted for the malto-oligosaccharides in the activity assay to test for substrate utilization of the glucoamylase (SEQ ID NO:594). Rate of glucose release for various substrates tested in 5 mM acetate buffer, 70°C. Substrates tested: maltose, maltotriose, panose, maltotetraose, and maltoheptaose, were all at 3 mM. The assay was then plotted in Figure 8. Then enzyme (SEQ ID NO:594) was able to hydrolyze straight-chain (1,4 linkages) dextrins down to maltose with a higher rate for the longer dextrins. The enzyme (SEQ ID NO:594) demonstrated low activity on 1,6 linkages as demonstrated by the substrate panose.

EXAMPLE 17: Glucoamylase Activity Assay: BCA Reducing Ends Assay

The following example describes an exemplary method for determining if a polypeptide is within the scope of the invention, for example, by a BCA reducing ends assay. Glucoamylase activity can be determined using the following methodology.

1. Prepare 2 substrate solutions, as follows:

a) 2% soluble starch (potato) pH 8 solution by dissolving 2 gm potato starch in 100 ml 100 mM sodium phosphate pH 8).

b) 2% soluble starch (potato) pH 10 solution by dissolving 2 gm potato starch in 100 ml 100 mM sodium carbonate.

Heat both solutions in a boiling water bath, while mixing, for 30-40 minutes until starch dissolves.

2. Prepare Solution A from 64 mg/ml sodium carbonate monohydrate, 24 mg/ml sodium bicarbonate and 1.95 mg/ml BCA (4,4'-dicarboxy-2,2'- biquinoline disodium salt (Sigma Chemical cat # D-8284). Added above to dH_2O.

3. Prepare solution B by combining 1.24 mg/ml cupric sulfate pentahydrate and 1.26 mg/ml L-serine. Add mixture to dH_2O.

4. Prepare a working reagent of a 1:1 ration of solutions A and B.

5. Prepare a Maltose standard solution of 10 mM Maltose in dH_2O, where the 10 mM maltose is combined in 2% soluble starch at desired pH to a final concentration of 0, 100, 200, 300, 400, 600 µM. The standard curve will be generated for each set of time-points. Since the curve is determined by adding 10 ul of the standards to the working reagent it works out to 0, 1, 2, 3, 4, 6 nmole maltose.

6. Aliquot 1 ml of substrate solution into microcentrifuge tubes, equilibrate to desired temperature (5 min) in heat block or heated water bath. Add 50 ul of enzyme solution to the inside of the tube lid.

7. While solution is equilibrating mix 5 ml of both solution A & B. Aliquot 100 ul to 96 well PCR plate. Set plate on ice.

8. After 5 minute temperature equilibration, close lid on tubes, invert and vortex 3 times. Immediately aliquot 10 ul into plate as t=0 (zero time point). Leave enzyme mixture in heat block and aliquot 10 ul at each desired time point (e.g. 0, 5, 10,15, 20, 30 minutes).

9. Ensure that 12 wells are left empty (only working reagent aliquotted) for the addition of 10 ul of standards, for the standard curve.

10. When all time points are collected and standards are added, cover plate and heated to 80° C for 35 min. Cool plate on ice for 10 min. Add 100 ul H2O to all wells. Mix and aliquot 100 ul into flat bottomed 96-well plate and read absorbance at 560 nm.

11. Zero each sample's time points against its own t=0 (subtract the average t=0 A560 value from other average A560 values). Convert the $A560_{(experimental)}$ to umole (Divide $A560_{(experimental)}$ by the slope of the standard curve (A560/umole). Generate a slope of the time points and the umole (in umole/min), multiply by 100 (as the umole value only accounts for the 10 ul used in the assay, not the amount made in the 1ml rxn). To get the specific activity divide the slope (in umole/min) by the mg of protein. All points should be done at a minimum in duplicate with three being best. Divide protein concentration (mg/ml) by any dilution to get mg used in assay. Divide the above slope by mg used in assay to get specific activity. See for example, Wong (2000) J. Agric. Food Chem. 48:4540-4543; Fox (1991) Anal. Biochem. 195, 93-96.

EXAMPLE 18: Screening for Glucoamylase activity

The following example describes an exemplary method for determining if a polypeptide is within the scope of the invention. Glucoamylase activity of clones can be assessed by a number of methods known in the art. The following is the general methodology that can be used.

The number of plaques screened, per plate, can be approximately 10,000 pfu's. For each DNA library: about 50,000 plaques per isolated library and 200,000 plaques per non-isolated library can be screened depending upon the pfu titer for the λ Zap Express amplified lysate.

Titer determination of Lambda Library

8) µL of Lambda Zap Express amplified library stock added to 600µL *E. coli* MRF' cells (OD_{600}=1.0). To dilute MRF' stock, 10mM $MgSO_4$ is used.
9) Incubate at 37°C for 15 minutes.
10) Transfer suspension to 5-6mL of NZY top agar at 50 °C and gently mix.
11) Immediately pour agar solution onto large (150mm) NZY media plate.
12) Allow top agar to solidify completely (approximately 30 minutes), then invert plate.

13) Incubate the plate at 39 °C for 8-12 hours.

14) Number of plaques is approximated. Phage titer determined to give 10,000 pfu/plate. Dilute an aliquot of Library phage with SM buffer if needed.

Substrate screening

13) Lambda Zap Express (50,000 pfu) from amplified library added to 600µL of *E. coli* MRF' cells (OD600=1.0). For non-environment libraries, prepare 4 tubes (50,000 pfu per tube).

14) Incubate at 37 °C for 15 minutes.

15) While phage/cell suspension are incubating, 1.0mL of red starch substrate (1.2% w/v) is added to 6.0mL NZY top agar at 50 °C and mixed thoroughly. Keep solution at 50°C until needed.

16) Transfer 1/5 (10,000 pfu) of the cell suspension to substrate/top agar solution and gently mixed.

17) Solution is immediately poured onto large (150mm) NZY media plate.

18) Allow top agar to solidify completely (approximately 30 minutes), then invert plate.

19) Repeat procedures 4-6 four times for the rest of the cell suspension (1/5 of the suspension each time).

20) Incubate plates at 39°C for 8-12 hours.

21) Plate observed for clearing zones (halos) around plaques.

22) Plaques with halos are cored out of agar and transferred to a sterile micro tube. A large bore 200µL pipette tip works well to remove (core) the agar plug containing the desired plaque.

23) Phages are re-suspended in 500µL SM buffer. 20µL Chloroform is added to inhibit any further cell growth.

24) Pure phage suspension is incubated at room temperature for 4 hours or overnight before next step.

Isolation of pure clones

12) 10µL of re-suspended phage suspension is added to 500µL of *E. coli* MRF' cells (OD600=1.0).

13) Incubate at 37°C for 15 minutes.

14) While phage/cell suspension is incubating, 1mL of red starch substrate (1.2% w/v) is added to 6.0mL NZY top agar at 50 °C and mixed thoroughly. Keep solution at 50 °C until needed.

15) Cell suspension is transferred to substrate/top agar solution and gently mixed.

16) Solution is immediately poured onto large (150mm) NZY media plate.

17) Allow top agar to solidify completely (approximately 30 minutes), then invert plate.

18) Plate incubated at 39°C for 8-12 hours.

19) Plate observed for a clearing zone (halo) around a single plaque (pure clone). If a single plaque cannot be isolated, adjust titer and re-plate phage suspension.

20) Single plaque with halo is cored out of agar and transferred to a sterile micro tube. A large bore 200µL pipette tip works well to remove (core) the agar plug containing the desired plaque. To amplify the titer, core 5 single active plaques into a micro tube.

21) Phages are re-suspended in 500µL SM buffer. 20µL Chloroform is added to inhibit any further cell growth.

22) Pure phage suspension is incubated at room temperature for 4 hours or overnight before next step. The pure phage suspension is stored at –80 °C by adding DMSO into the phage suspension (7% v/v).

Excision of pure clone

17) 100µL of pure phage suspension is added to 200µL *E. coli* MRF' cells (OD600=1.0). To this, 1.0µL of ExAssist helper phage (>1 x 106 pfu/mL; Stratagene) is added. Use 2059 Falcon tube for excision.

18) Suspension is incubated at 37°C for 15 minutes.

19) 3.0 mL of 2 x YT media is added to cell suspension.

20) Incubate at 30 °C for at least 6 hours or overnight while shaking.

21) Tube transferred to 70°C for 20 minutes. The phagemid suspension can be stored at 4°C for 1 to 2 months.

22) 100 µL of phagemid suspension transferred to a micro tube containing 200µL of *E. coli* Exp 505 cells (OD600=1.0).

23) Suspension incubated at 37 °C for 15 minutes.

24) 300µL of SOB is added to the suspension.

25) Suspension is incubated at 37°C for 30 to 45 minutes.

26) 100µL of suspension is transferred to a small (90mm) LB media plate containing Kanamycin (LB media with Kanamycin 50µg/mL) for Zap Express DNA libraries or Ampicillin (LB media with Kanamycin 100µg/mL) for Zap II DNA libraries.

27) The rest of suspension is transferred to another small LB media plate.

28) Use sterile glass beads to evenly distribute suspension on the plate.

29) Plates are incubated at 30°C for 12 to 24 hours.

30) Plate observed for colonies.

31) Inoculate single colony into LB liquid media containing suitable antibiotic and incubate at 30 °C for 12 to 24 hours.

32) Glycerol stock can be prepared by adding 80% glycerol into liquid culture (15% v/v) and stored at –80 °C.

Activity verification

7) 50µL of liquid culture is transferred to a micro tube. Add 500µL of 8% pH7 Amylopectin Azure into the same tube. Prepare 2 tubes for each clone.

8) Activity is tested at 50°C for 3 hours and overnight. Use pH 7 buffer as control.

9) Cool the test specimen at ice-water bath for 5 minutes.

10) Add 750µL of Ethanol and mixed thoroughly.

11) Centrifuge at 13000 rpm (16000 g's) for 5 minutes.

12) Measure OD of the supernatant at 595nm.

RFLP analysis

13) 1.0mL of liquid culture is transferred to a sterile micro tube.

14) Centrifuge at 13200 rpm (16000 g's) for 1 minute.

15) Discard the supernatant. Add another 1.0 mL of liquid culture into the same sterile micro tube.

16) Centrifuge at 13200 rpm (16000 g's) for 1 minute.

17) Discard the supernatant.

18) Follow QIAprep spin mini kit protocol for plasmid isolation.

19) Check DNA concentration using BioPhotometer.

20) Use Sac I and Kpn I for first double digestion. Incubate at 37 °C for 1 hour.

21) Use Pst I and Xho I for second double digestion. Incubate at 37 °C for 1 hour.

22) Add Loading dye into the digested sample.

23) Run the digested sample on a 1.0% agarose gel for 1-1.5 hours at 120 volts.

24) View gel with gel imager. All clones with a different digest pattern will be sent for sequence analysis.

EXAMPLE 19: Assay for glucoamylases

The following example describes an exemplary method for determining if a polypeptide is within the scope of the invention.

Preparation Of Host Cultures

5. Start an overnight culture of XL1-Blue MRF' host cells. Use a single colony from a streak plate to inoculate 10 mL LB supplemented with 20 ug/mL tetracycline. Grow overnight culture shaking at 37°C for at least 16 hours.
6. Using aseptic technique, inoculate a fresh 100 mL of LB_{Tet} day culture with XL1-Blue MRF' host from the overnight LB_{Tet} culture.
7. Grow in a 37°C shaker until the OD reaches 0.75 – 1.0.
8. Pellet host cells at 1000 x g for 10 minutes and gently resuspend in 10 mM $MgSO_4$ at OD5.
9. Dilute a small amount of host cells to OD1 for use in titering and pintooling.
10. Host preparations can be used for up to 1 week when stored on ice or at 4°C.

-To shorten growth time for the day culture, use ½X the usual Tet concentration in LB (½X = 10 ug/mL), or omit the antibiotic altogether.

-Do not use NZY when selecting with Tetracycline. The high Mg^{++} concentration in NZY medium renders Tet inactive.

Titering Lambda Libraries

11. Place three sterile microfuge tubes in a rack.
12. Aliquot 995 uL prepared host cells in one tube and 45 uL prepared OD1 host cells into each of the two remaining tubes.
13. Add 5 uL of lambda library to the tube containing 995 uL host cells and mix by vortexing. This results in a dilution factor of 200.
14. Prepare 1/2,000 and 1/20,000 dilutions by consecutively adding 5 uL of previous dilution to the remaining two tubes containing 45 uL prepared host cells. Mix by vortexing after each dilution was made.
15. Allow phage to adsorb to host by incubating at 37°C for 15 minutes.

16. Meanwhile, pipet 100 uL of prepared OD1 host cells to each of three Falcon 2059 tubes.
17. Add 5 uL of each dilution to a separate 2059 tube containing host cells.
18. Plate each by adding 3 mL top agar to each tube and quickly pour over 90 mm NZY plates. Ensure a smooth, even distribution before the top agar hardens.
19. Invert plates and incubate at 37°C overnight.
20. Count plaques and calculate titer of the library stock (in plaque forming units (pfu) per uL).

Lambda Microtiter Screening For glucoamylases

Preparation

5. Prepare a sufficient amount of XL1-Blue MRF' host culture, as described above, for the amount of screening planned. A culture of 100 mL is usually sufficient for screening 2-3 libraries.
6. Autoclave several bottles compatible with the QFill2 dispenser. These are the wide-mouth Corning bottles, 250 mL containing a sealing ring around the lip.
7. Make sure there are sufficient amounts of plates, top agar, BODIPY starch, red starch solution, etc. available for the screen.
8. Schedule the Day 2 robot run with a representative from Automation.

Day 1

10. Label the 1536-well plates (black) with library screen and plate number. Tough-Tags™ tube stickers, cut in half width-wise, are ideal for labeling 1536 well plates.
11. Calculate volumes of library, host cells and NZY medium necessary for the screen. This is easily done with an Excel spreadsheet.
12. Combine the calculated volumes of lambda library and OD5 host cells in a sterile 250 mL wide-mouth Corning bottle (containing a sealing ring).
13. Allow adsorption to occur at 37°C for 15 minutes.
14. Add the calculated volume of NZY medium and mix well. This is referred to as the cell-phage-medium suspension.
15. Perform a concomitant titer by combining 50 uL of the cell-phage-medium suspension with 250 uL of OD1 host cells in a Falcon 2059 tube, then plating with 9 mL of top agar onto a 150 mm NZY plate. Incubate concomitant titer plate at 37°C overnight.

16. Load the dispenser with the remainder of the suspension and array each labeled 1536-well plate at 4 uL per well. If the dispenser leaves air bubbles in some wells, they can be removed by centrifuging the plates at 200 x g for 1 minute.

17. Add 0.5 uL of positive control phage to well position AD46 of at least two of the assay plates. Use a strong glucoamylase-positive lambda clone for this purpose. The lambda versions of SEQ ID NO.: 113 or SEQ ID NO.: 199 are good choices for positive controls.

18. Incubate assay plates at 37°C overnight in a humidified (≥95%) incubator.

Day 2

21. Count the pfu on the concomitant titer plate and determine the average seed density per well (in pfu per well).

22. Pintool at least 2 plates of each library screen (preferably the 2 containing positive controls) as follows:

a) Prepare 2 host lawn plates to act as a surface on which to pintool: combine 250 uL of OD1 host cells with 2 mL 2% red starch and plate with 9 mL top agar onto 150 mm NZY plates. Hold each plate as level as possible as the top agar solidifies in order to produce an even hue of red across the plate.

b) Using a twice flame-sterilized 1536 position pintool, replicate at least 2 of the screening plates onto the host lawn plates.

c) Place the pintooled recipient plates in a laminar flow hood with the lids off for about 15-30 minutes (to vent off excess moisture).

d) Replace the lids and incubate inverted at 37°C overnight.

23. Prepare the 2X BODIPY starch substrate buffer as follows:

a) Calculate the total volume of 2X substrate buffer solution needed for all screening plates at 4 uL per well (including any extra deadspace volume required by the dispenser) and measure this amount of 100 mM CAPS pH 10.4 into a vessel appropriate for the dispenser used.

b) Retrieve enough 0.5 mg tubes of BODIPY starch to produce the required volume of 2X substrate buffer [calculated in step a) above] at a final concentration of 20-30 ug/mL.

c) Dissolve each 0.5 mg tube in 50 uL DMSO at room temperature, protected from light, with frequent vortexing. This takes more than 15 minutes; some production lots of BODIPY starch dissolve better than others.

d) Add 50 uL 100mM CAPS buffer pH 10.4 to each tube and mix by vortexing.

e) Pool the contents of all tubes and remove any undissolved aggregates by centrifuging for 1 minute at maximum speed in a microfuge.

f) Add the supernatant to the rest of the 100 mM CAPS buffer measured in step a) above.

g) Protect the 2X substrate buffer from light by wrapping in foil.

24. Take plates and substrate buffer to the automation room and program the robot with the following parameters:

a) dispense 4 uL substrate buffer per well

b) 1st read at 1 hour post-substrate, 2nd read at 9 hours, and third read at 17 hours; with 37°C incubation between reads

c) excitation filter: 485 nm; emission filter: 535 nm

d) set the Spectrafluor gain at 70, or the optimal gain for the batch of 2X substrate buffer prepared.

e) ensure that the incubator used will protect assay plates from light.

Day 3

4. Check pintooled plates for clearings in the bacterial lawn at all positions corresponding to wells on the associated assay plate. Also check for clearings in the red starch in any of the pin positions. If plates containing positive controls were used for pintooling, you should be able to see a large clearing zone in the red background. Be wary of contaminants that also form clearing zones in red starch (see comment "Contaminants That Form Clearing Zones in Red Starch").

5. Identify putative hits from the data file produced by the robot computer. The KANAL program produced by Engineering simplifies data analysis. As a rule of thumb, a putative hit is characterized as a well having signal intensity rising at least 1.5 fold over background.

6. For each putative, remove 2 uL from the well and add to a tube containing 500 uL SM buffer and 50 uL CHCl3. Vortex to mix and store at 4°C. This solution will be referred to hereafter as the 4e-3 stock. The original screening

plates should be stored at 4°C, protected from light, at least until breakouts are complete.

This is the recommended method of breaking out putative hits. It is a liquid phase assay that relies on confirmation of activity on BODIPY starch. Alternatively, putative hits can be plated directly onto solid phase plates containing red starch such that 2,000-3,000 pfu per hit are examined for clearing zones. However, inability to observe clearing zones on red starch is not necessarily an indication that a putative hit was a false positive. It would then need to be assayed using the format in which it was originally identified (i.e., liquid phase using BODIPY starch as substrate). In addition, very weak positives are more easily identified using the method detailed below.

Day 1

25. In a sterile 50 mL conical tube, combine 0.5 mL OD5 host cells with 45.5 mL NZY. This will be referred to as the host-medium suspension.
26. For each putative hit to be analyzed, aliquot 1 mL of host-medium suspension into each of 3 three sterile microfuge tubes.
27. Set the 12-channel pipetman in multidispense mode with an aliquot size of 20 uL and an aliquot number of **2x**. Mount the pipetman with a clean set of sterile tips.
28. Pour about 1 mL of host-medium suspension into a new sterile solution basin and load the multichannel pipetman.
29. Dispense 20 uL per well into the last row (row P) of a black 384-well plate (12 channels x 2 = 24 wells). This row will be used later for the controls.
30. Expel the remaining liquid in the tips by touching the tips against the surface of the basin and pressing the RESET button on the pipetman. Lay the pipetman down in a way to prevent contamination of the tips. There is no need to change the tips at this point.
31. Pour the remainder of the fluid in the basin into a waste container (like a beaker) taking care to avoid splash-back contamination.
32. For the first putative to be analyzed, take 111 uL of the 4e-3 stock (see Day 2 in *Lambda Microtiter Screening for glucoamylases*) and add it to the first in a set of three tubes containing 1 mL host-medium suspension (step 2). Vortex to mix. This is *Dilution A*.

33. Take 111 uL of Dilution A and add to the next tube in the set. Vortex to mix. This is *Dilution B.*
34. Take 111 uL of Dilution B and add to the last tube in the set. Vortex to mix. This is *Dilution C.* You should now have three dilutions of phage, where concentrations of each differ by a factor of 10.
35. Pour the contents of Dilution C (the most dilute of the 3 samples) into the solution basin and load the multichannel pipetman.
36. Dispense 20 uL per well into the first row of the 384-well plate (12 channels x 2 = 24 wells).
37. Expel the remaining liquid in the tips by touching the tips against the surface of the basin and pressing the RESET button on the pipetman. Lay the pipetman down in a way to prevent contamination of the tips. There is no need to change the tips at this point.
38. Empty the basin as described above.
39. Pour the contents of Dilution B into the same basin and load the multichannel pipetman.
40. Dispense 20 uL per well into the second row of the 384-well plate.
41. Perform steps 13-16 similarly to dispense Dilution A into the third row of the plate.
42. After all three dilutions have been arrayed into the first 3 rows of the plate, discard all tips and the solution basin into the biohazardous waste container.
43. Mount the pipetman with a clean set of sterile tips and open a new sterile solution basin.
44. Repeat steps 8-19 for each remaining putative hit, using remaining rows on the plate up to row O. Five putative hits can be analyzed on one 384-well plate, with the last row (row P) saved for the controls.
45. Add 0.5 uL of each control to a separate well. Use at least 2-3 separate controls, preferably covering a range of activity.
46. Incubate assay plates at 37°C overnight in a humidified (≥95%) incubator.

<u>Day 2</u>

47. Pintool all breakout plates onto a host lawn with red starch using the same method described for Day 2 in *Lambda Microtiter Screening for glucoamylases*, except that a 384 position pintool is used.

48. Prepare the 2X BODIPY starch substrate buffer as follows:
    a) Calculate the total volume of 2X substrate buffer solution needed for all breakout plates at 20 uL per well (including any extra deadspace volume required by the dispenser) and measure this amount of 100 mM CAPS pH 10.4 into a vessel appropriate for the dispenser used.
    b) Retrieve enough 0.5 mg tubes of BODIPY starch to produce the required volume of 2X substrate buffer [calculated in step a) above] at a final concentration of 20-30 ug/mL.
    c) Dissolve each 0.5 mg tube in 50 uL DMSO at room temperature, protected from light, with frequent vortexing. This takes more than 15 minutes; some production lots of BODIPY starch dissolve better than others.
    d) Add 50 uL 100mM CAPS buffer pH 10.4 to each tube and mix by vortexing.
    e) Pool the contents of all tubes and remove any undissolved aggregates by centrifuging for 1 minute at maximum speed in a microfuge.
    f) Add the supernatant to the rest of the 100 mM CAPS buffer measured in step a) above.
    g) Protect the 2X substrate buffer from light by wrapping in foil.
49. Dispense 20 uL per well into all breakout plates.
50. Wrap all plates in aluminum foil and incubate at room temperature for 2-6 hours.
51. Read each plate in the Spectrafluor with the following settings:
    a) fluorescence read (excitation filter: 485 nm; emission filter: 535 nm)
    b) plate definition: 384 well black
    c) read from the top
    d) optimal gain
    e) number of flashes: 3
52. On the resulting Excel spreadsheet, chart each putative's 3 rows in a separate graph and check for activity. Ensure that the positives controls produced signals over background.
53. For each putative that appears to have a real signal among the wells, harvest a sample from a positive well as follows:
    a) Select a positive well from a row representing the highest initial dilution.
    b) Transfer 2 uL from that well into a tube containing 500 uL SM and 50 uL $CHCl_3$. This is referred to as the breakout stock.

c) Store at 4°C.

54. Using methods previously described, plate about 10 uL of each breakout stock onto 150 mm NZY plates using red starch. The objective is to obtain several (at least 20) well-separated plaques from which to core isolates.

Day 3

55. Check pintooled plates for an acceptable incidence of clearings in the bacterial lawn corresponding to wells on the associated assay plate. Also check for clearings in the red starch in the positive controls and in any tested putatives. Be wary of contaminants that also form clearing zones in red starch (see below).

56. From the solid phase plates containing dilutions of breakout stocks, core several isolated plaques, each into 500 uL SM with 50 uL $CHCl_3$. This is referred to as the isolate stock.

57. The isolate stocks can then be individually tested on BODIPY starch using methods described above. This step can be skipped if the plaque that was cored in step 2 produced a clearing zone in the red starch background. The isolate stocks were then be individually tested on BODIPY starch using methods described above. However, this step may be skipped if the plaque that was cored in step 2 produced a clearing zone in the red starch background.

Excisions

Day 1

58. In a Falcon 2059 tube, mix 200 uL OD1 XL1-Blue MRF' host, 100 uL lambda isolate stock and 1 uL ExAssist phage stock.

59. Incubate in 37°C shaker for 15 minutes.

60. Add 3 mL NZY medium.

61. Incubate in 30°C shaker overnight.

Day 2

10. Heat to excision tube to 70°C for 20 minutes.

11. Centrifuge 1000 x g for 10 minutes.

12. In a Falcon 2059 tube, combine 50 uL supernatant with 200 uL EXP505 OD1 host.

13. Incubate in 37°C shaker for 15 minutes.

14. Add 300 uL SOB medium.

15. Incubate in 37C shaker for 30-45 minutes.
16. Plate 50 uL on large LB_{Kan50} plate using sterile glass beads. If the plates are "dry", extra SOB medium can be added to help disburse the cells.
17. Incubate plate at 30°C for at least 24 hours.
18. Culture an isolate for sequencing and/or RFLP.

Growth at 30°C reduces plasmid copy number and is used to mitigate the apparent toxicity of some glucoamylase clones.

Contaminants That Form Clearing Zones in Red Starch

When using red starch on solid medium to assay phage for glucoamylase activity, it is common to see contaminating colony forming units (cfu) that form clearing zones in the red starch. For pintooled plates, it is important to distinguish glucoamylase-positive phage clones from these contaminants whenever they align with a particular well position. The source of the contaminating microbes is presumably the 2% red starch stock solution, which cannot be sterilized by autoclaving or by filtering after preparation. It is thought that they are opportunistic organisms that survive by metabolizing the red starch. In order to reduce these contaminants, use sterile technique when making 2% red starch solutions and store the stocks either at 4°C or on ice.

Assay Using RBB-starch

75µl of RBB-starch substrate (1% RBB-insoluble corn starch in 50mM NaAc buffer, pH=4.5) can be added into each well of a new 96-well plate (V-bottom). Five micro-liters of enzyme lysate can be transferred into each well with substrate using Biomek or Zymark. The plates can be sealed with aluminum sealing tape and shaken briefly on the shaker. The plates can be incubated at 90^0C for 30 minutes, followed by cooling at room temperature for about 5 to 10 minutes. One hundred micro-liters of 100% ethanol is added to each well, the plates sealed and shaken briefly on the shaker. The plates are then centrifuged 4000rpm for 20 minutes using bench-top centrifuge. 100µl of the supernatant is transferred into a new 96-well plate (flat bottom) by Biomek and read OD_{595}.

Assay using FITC-starch

Add 50µl of substrate (0.01% FITC-starch in 100mM NaAc buffer, pH=4.5) into each well of a new 384-well plate. Transfer 5µl of enzyme lysate into each well with substrate and incubate the plate at room temperature overnight. The polarization change of the substrate, excitation 485nm, emission 535nm, is read for each well. 96 well plates can be used for all assays.

Example 20: Exemplary protocol for liquefying starch and measuring results

The following example described and exemplary protocol for liquefying starch. Reaction Conditions: 100 mM PO_4 pH 6.5, 1% (w/w) liquefied starch DE 12 at 55°C. Both TLC and HPLC assays can be done to verify activity.

An exemplary protocol for the saccharification of liquefied starch at pH 6.5:

- Adjust the pH of the liquefied starch to the pH at which the saccharification(s) will be performed. Liquefy starch in 100 mM sodium acetate buffer, pH 4.5 with 100 mM sodium phosphate salts added so that before saccharification, the pH could be adjusted to pH 6.5.
- Weigh 5 gram samples of liquefied starch into tared bottles.
- Use 0.04% (w/w) Optidex L-400 or approximately 400 mL of 1-10 diluted stock Optidex L-400 per 100 grams of liquefied starch.
- Calculate the milligrams of Optidex L-400 contained in the 400 mL of 1-10 diluted stock Optidex L-400. Next, calculate the volume of lysates needed to give the same concentration of enzyme as the Optidex L-400.
- Add enzymes to liquefied starch samples and incubate at desired temperature (50C°). After 18 hours determine DE and prepare a sample for HPLC analysis.

An exemplary DE Determination:

Exemplary Neocuproine Assay:

A 100ml sample can be added to 2.0ml of neocuproine solution A (40g/L sodium carbonate, 16g/L glycine, 0.45g/L copper sulfate). To this can be added 2.0 ml of neocuproine solution B (1.2g/L neocuproine hydrochloride-Sigma N-1626). The tubes can be mixed and heated in a boiling water bath for 12 minutes; cooled, diluted to 10ml volume with DI water and the OD read at 450nm on the spectrophotometer. The glucose

equivalent in the sample can be extrapolated from the response of a 0.2mg/ml glucose standard run simultaneously.

Exemplary HPLC Analysis:

Saccharification carbohydrate profiles are measured by HPLC (Bio-Rad Aminex HPX-87A column in silver form, 80°C) using refractive index detection. Mobile phase is filtered Millipore water used at a flow rate of 0.7 ml/min. Saccharification samples are diluted 1-10 with acidified DI water (5 drops of 6 M HCl into 200 mL DI water) then filtered through a 0.45 mm syringe filter. Injection volume is 20 uL.

Exemplary TLC:

Reaction products can be w/d at hourly timepoints and spotted and dried on a TLC plate. The plate can be then developed in 10:90 water:isopropanol and visualized with either a vanillin stain or CAM stain and then heated to show reducible sugars. The liquefied starch can be partially hydrolyzed to glucose in cases where activity was observed.

EXAMPLE 21: Starch Liquefaction using glucoamylases

This example describes an exemplary method of the invention for liquefying starch using glucoamylases of the invention. Glucoamylase concentrate can be prepared from fermentation broths by heat treatment, cell washing, alkaline extraction using microfiltration and ultrafiltration (48% overall yield). The UF concentrate can be neutralized with acetic acid and formulated with 30% glycerol at pH 4.5. The activity level of a commercial product can be about 120 U^1/g – 0.5 kg/ ton starch.

Exemplary glucoamylase activity assay

A 1 mL cuvette containing 950 µL of 50 mM MOPS pH 7.0 containing 5 mM PNP-α- D—hexa-(1→4)-glucopyranoside is placed in the Peltier temperature controller of the Beckman DU-7400 spectrophotometer preheated to 80°C. The spectrophotometer is blanked at 405nm and 50 µL of the enzyme solution is added to the cuvette, mixed well and the increase in the OD_{405nm} is monitored over a one-minute interval. The $\Delta OD_{405nm/min}$ rate is converted to a standard unit of µmole/minute from the OD_{405nm} response of 50 µL of 1 µmole/mL PNP in 950 mL 50 mM MOPS at pH 7.0 - 80°C. One standard unit of thermostable alpha glucoamylase (DTAA) is equal to the

amount of enzyme that will catalyze the release of 1 μmole/mL/minute of pNP under the defined conditions of the assay.

Standard Glucoamylase Activity Assay

A 1 mL cuvette containing 950 μL of 50 mM MOPS pH 7.0 containing 5 mM pNP-α- D-glucopyranoside is placed in the Peltier temperature controller of the Beckman DU-7400 spectrophotometer preheated to 60°C. The spectrophotometer is blanked at 405nm and 50 μL of the enzyme solution is added to the cuvette, mixed well and the increase in the OD_{405nm} is monitored over a one-minute interval. The ΔOD_{405nm}/min rate is converted to a standard unit of μmole/minute from the OD_{405nm} response of 50 μL of 1 μmole/mL pNP in 950 mL 50 mM MOPS at pH 7.0-60°C. One standard Diversa unit of glucoamylase (DGA) is equal to the amount of enzyme that will catalyze the release of 1 μmole/mL/minute of pNP under the defined conditions of the assay.

Dextrose Equivalent Determination

The neocuproine method is used to measure the DE. Selected samples were measured by both the Invention procedure and by a GPC analyst using the GPC Fehlings procedure.

Neocuproine Assay

A 100 μl sample is added to 2.0 ml of neocuproine solution A (40 g/L sodium carbonate, 16g/L glycine, 0.45g/L copper sulfate). To this is added 2.0 ml of neocuproine solution B (1.2 g/L neocuproine hydrochloride-Sigma N-1626). The tubes were mixed and heated in a boiling water bath for 12 minutes; cooled, diluted to 10ml volume with DI water and the OD read at 450 nm on the spectrophotometer. The glucose equivalent in the sample is extrapolated from the response of a 0.2mg/ml glucose standard run simultaneously.

The starch sample is diluted ~1 to 16 with DI water with the exact dilution recorded. Ten milliliters of the diluted sample is added to 20 mls of DI water. Ten milliliters of Fehlings solution A and B were added to the diluted starch. The sample is boiled for 3 minutes and cooled on ice. Ten milliliters of 30% KI and 10ml of 6N H_2SO_4 is added. The solution is titrated against 0.1N sodium thiosulfate. The titrant volume is recorded and used to calculate the DE.

Residual Starch Determination

Post-saccharification samples were checked for residual starch using the Staley iodine procedure.

Twenty grams of sample is weighed into a large weigh dish. 45 µL of Iodine solution is added to the weigh dish and the starch solution is mixed well. Dark blue indicates the presence of starch, a light blue-green indicates slight starch, light green indicates a trace of starch and yellow-red, absence of starch. Iodine solution is prepared by dissolving 21.25 grams of iodine and 40.0 grams of potassium iodide in one liter of water.

Oligosaccharide Profile

Liquefaction and saccharification carbohydrate profiles were measured by HPLC (Bio-Rad Aminex HPX-87C column in calcium form – 80°C) using refractive index detection.

Gel Permeation Chromatography

The molecular weight distribution is determined by chromatography on a PL Aquagel-OH column with mass detection by refractive index (Waters Model 2410). A Viscotek Model T60 detector is used for continuous viscosity and light scattering measurements.

Capillary Electrophoresis

Beckman Coulter P/ACE MDQ Glycoprotein System – separation of APTS derivatized oligosaccharides on a fused silica capillary - detection by laser-induced fluorescence.

Primary Liquefaction

Line starch directly from the GPC process is pumped into a 60 liter feed tank where pH, DS (dry solids) and calcium level can be adjusted before liquefaction. The glucoamylase is added to the slurry. The 32% DS slurry is pumped at 0.7 liter/minute by a positive displacement pump to the jet - a pressurized mixing chamber where the starch slurry is instantaneously heated to greater than 100°C by steam injection. The gelatinized partially liquefied starch is pumped through a network of piping (still under pressure) to give the desired dwell time (5 minutes) at temperature. The pressure is released into a flash tank and samples can be taken. Samples were taken in duplicate.

Secondary Liquefaction

The liquefied starch is collected in one liter glass bottles and held in a water bath at 95°C for 90 minutes.

Saccharification

Liquefied starch is cooled to 60°C, the pH adjusted to 4.5 and the samples treated with glucoamylase. Saccharification progress is monitored over time by HPLC.

Saccharification

The liquefied syrups produced with each glucoamylase were adjusted to approximately pH 2.5 with 6N HCl immediately after the 90 minute secondary liquefaction to inactivate any residual glucoamylase. The syrups were then adjusted to pH 4.5, placed in a 60°C water bath and saccharified with three levels of glucoamylase. The extent of saccharification is monitored by HPLC at 18 to 88 hour time points.

The liquefied syrups were saccharified with the standard dosage – 0.04% of a double-strength glucoamylase - and two lower dosages (50% and 25%) to monitor any differences in the saccharification progress.

Saccharification Progress - % dextrose development vs time – 0.04% glucoamylase.

EXAMPLE 22: Starch Liquefaction at pH 4.5 using glucoamylases

The conversion of starch to glucose can be catalyzed by the sequence action of two enzymes: amylases (e.g., alpha-amylases), including enzymes of the invention, to liquefy the starch (e.g., the hydrolysis of high molecular weight glucose polymers to oligosaccharides consisting of 2 to 20 glycose units, typically a dextrose equivalent of 10 to 12, by a glucoamylase of the invention), followed by saccharification with a glucoamylase (which can be a glucoamylase of the invention, e.g., SEQ ID NO:594). In one aspect, processing is in a corn wet milling plant producing a starch slurry having a pH or about 4.0 to 4.5. In one aspect, the pH is raised, e.g., to 5.8 to 6.0 before liquefaction to accommodate a glucoamylase with a low pH activity and stability. In one aspect, glucoamylases of the invention can liquefy starch at pH 4.5 to dextrose equivalents ranging from 12 to 18; in one aspect, using glucoamylases of the invention at levels of about 3 to 6 grams per ton of starch. In this aspect, use of glucoamylases of the

invention enables starch liquefaction to be conducted at pH 4.5.

In one aspect, starch liquefaction is conducted at pH 4.5 for 5 minutes at 105°C to 90 minutes at 95°C using glucoamylases of the invention. The quantity of enzyme is adjusted in order to adjust a target DE of 12 to 15 after liquefaction. In one aspect, the liquefied starch is then saccharified with a glucoamylase, e.g., an *Aspergillis* glucoamylase, for about 48 hours at about pH 4.5 and 60°C. If the saccharified syrup did not contain at least 95% glucose, the target liquefaction DE is raised and the saccharification repeated until the liquefaction eventually did produce a saccharified syrup containing more than 95% glucose. The glucoamylase protein required to produce a suitable liquefied feedstock for saccharification is determined by PAGE.

A number of embodiments of the invention have been described. Nevertheless, it will be understood that various modifications may be made without departing from the spirit and scope of the invention. Accordingly, other embodiments are within the scope of the following claims.

2013201807 09 Sep 2015

THE CLAIMS DEFINING THE INVENTION ARE AS FOLLOWS:

1. An isolated, synthetic, or recombinant nucleic acid comprising:

(a) a nucleic acid sequence having at least 85% sequence identity to the sequence SEQ ID NO:605, wherein the nucleic acid encodes at least one polypeptide having an alpha-amylase activity; or

(b) a nucleic acid sequence encoding a polypeptide having at least 85% sequence identity to the sequence SEQ ID NO:606, or an enzymatically active fragment thereof having alpha-amylase activity; or

(c) a nucleic acid sequence encoding a polypeptide having at least 85% sequence identity to the sequence SEQ ID NO: 606, that hybridizes under stringent conditions to a nucleic acid comprising SEQ ID NO:605, wherein the nucleic acid encodes a polypeptide having an alpha amylase activity, and the stringent conditions comprise a wash step comprising a wash in 0.2X SSC at a temperature of about 65°C for about 15 minutes; or

(d) the nucleic acid sequence SEQ ID NO:605, wherein the nucleic acid encodes at least one polypeptide having an alpha-amylase activity; or

(e) a nucleic acid sequence of any one of (a) to (d) encoding an alpha-amylase lacking a signal sequence; or

(f) a nucleic acid sequence of any one of (a) to (e) encoding an alpha-amylase having a heterologous sequence or signal sequence; or

(g) a fully complementary nucleic acid sequence to any one of (a) to (f).

2. An expression cassette, a vector, or a cloning vehicle comprising a nucleic acid:

(a) comprising a sequence according to claim 1; or

(b) comprising a sequence according to claim 1, wherein the cloning vehicle comprises a viral vector, a plasmid, a phage, a phagemid, a cosmid, a fosmid, a bacteriophage or an artificial chromosome; or

(c) comprising a sequence according to claim 1, wherein the viral vector comprises an adenovirus vector, a retroviral vector or an adeno-associated viral vector; or

2013201807 09 Sep 2015

(d) comprising a sequence according to claim 1, comprising a bacterial artificial chromosome (BAC), a plasmid, a bacteriophage P1-derived vector (PAC), a yeast artificial chromosome (YAC), or a mammalian artificial chromosome (MAC).

3. A transformed cell:

(a) comprising a heterologous nucleic acid that comprises a sequence according to claim 1, or an expression cassette, vector, or cloning vehicle according to claim 2; or

(b) comprising a heterologous nucleic acid that comprises a sequence according to claim 1, or an expression cassette, vector, or cloning vehicle according to claim 2, wherein the cell is a bacterial cell, a mammalian cell, a fungal cell, a yeast cell, an insect cell or a plant cell.

4. An isolated, synthetic, or recombinant polypeptide having an alpha-amylase activity comprising:

(a) an amino acid sequence having at least 85% sequence identity to the sequence SEQ ID NO:606, or an enzymatically active fragment thereof having an alpha-amylase activity, wherein the fragment is 40 or more amino acid residues in length; or

(b) an amino acid sequence encoded by a nucleic acid according to claim 1; or

(c) an amino acid sequence according to (a) or (b) lacking a signal sequence; or

(d) an amino acid sequence according to (a), (b) or (c) further comprising a heterologous sequence or heterologous signal sequence; or

(e) an amino acid sequence according to any one of (a) to (d) further comprising a polysaccharide, or wherein the polypeptide comprises at least one glycosylation site.

5. A protein preparation comprising a polypeptide of claim 4, wherein the protein preparation comprises a liquid, a solid, or a gel.

6. A method for hydrolyzing a starch, or removing, or liquefying the starch, comprising the following steps:

2013201807 09 Sep 2015

(a) providing a polypeptide having an alpha-amylase activity, wherein the polypeptide comprises a polypeptide of claim 4, or a polypeptide encoded by a nucleic acid of claim 1;

(b) providing a composition comprising a starch; and

(c) contacting the polypeptide of step (a) with the composition of step (b) under conditions wherein the polypeptide hydrolyzes the starch, or removes or liquefies the starch.

7. A detergent composition comprising a polypeptide of claim 4, or a polypeptide encoded by a nucleic acid of claim 1.

8. A food supplement, feed supplement, feed or a food comprising a polypeptide of claim 4, or a polypeptide encoded by a nucleic acid of claim 1.

9. A composition comprising a polypeptide of claim 4, or a polypeptide encoded by a nucleic acid of claim 1.

10. A method for textile processing or desizing comprising the following steps:

(a) providing a polypeptide having an alpha-amylase activity, wherein the polypeptide comprises a polypeptide of claim 4, or a polypeptide encoded by a nucleic acid of claim 1;

(b) providing a textile; and

(c) contacting the polypeptide of step (a) and the textile of step (b) under conditions wherein the alpha-amylase can process or desize the textile.

11. A method for treating a paper, a paper product, a paper pulp or a fiber comprising the following steps:

(a) providing a polypeptide having an alpha-amylase activity, wherein the polypeptide comprises a polypeptide of claim 4, or a polypeptide encoded by a nucleic acid of claim 1;

(b) providing a composition comprising a paper, a paper product, a paper pulp or a fiber; and

2013201807 09 Sep 2015

(c) contacting the polypeptide of step (a) and the composition of step (b) under conditions wherein the polypeptide can treat the paper, paper product, paper pulp or fiber.

12. A method for producing a high-maltose or a high-glucose syrup comprising the following steps:

(a) providing a polypeptide having an alpha-amylase activity, wherein the polypeptide comprises a polypeptide of claim 4, or a polypeptide encoded by a nucleic acid of claim 1;

(b) providing a composition comprising a starch; and

(c) contacting the polypeptide of step (a) and the fabric of step (b) under conditions wherein the polypeptide of step (a) can hydrolyze the composition of step (b), thereby producing a high-maltose or a high-glucose syrup.

13. A method for improving the flow of the starch-containing production fluids, comprising the following steps:

(a) providing a polypeptide having an alpha-amylase activity, wherein the polypeptide comprises a polypeptide of claim 4, or a polypeptide encoded by a nucleic acid of claim 1;

(b) providing a production fluid comprising a polysaccharide; and

(c) contacting the polypeptide of step (a) and the production fluid of step (b) under conditions wherein the alpha-amylase can hydrolyze the polysaccharide in the production fluid, thereby improving its flow by decreasing its density.

14. A method for using amylase in brewing or alcohol production comprising the following steps:

(a) providing a polypeptide comprising a polypeptide of claim 4, or a polypeptide encoded by a nucleic acid of claim 1;

(b) providing a composition used for brewing or in alcohol production comprising a starch; and

2013201807 09 Sep 2015

(c) combining the polypeptide of step (a) with the composition of the step (b) under conditions wherein the polypeptide can hydrolyze the starch in the composition used for brewing or alcohol production.

15. An oil well drilling fluid comprising a polypeptide of claim 4 or a polypeptide encoded by a nucleic acid of claim 1.

16. A method for changing the viscosity of a composition comprising treating the composition with a polypeptide of claim 4.

17. A method for aiding in the carrying away of drilling mud comprising treating the drilling mud with a composition comprising a polypeptide of claim 4 or a polypeptide encoded by a nucleic acid of claim 1.

18. A method for bio-bleaching a composition comprising treating the composition with a polypeptide of claim 4 or a polypeptide encoded by a nucleic acid of claim 1.

19. A method for making an ethanol-based fuel comprising the following steps:

(a) providing an alpha-amylase enzyme of claim 4;

(b) providing a composition comprising a starch; and

(c) contacting the amylase of (a) with the composition of (b) under conditions wherein the amylase hydrolyzes the starch.

20. A method for producing a food or feed comprising a recombinant alpha-amylase, the method comprising the steps of:

(a) providing a polypeptide having an alpha-amylase activity, wherein the polypeptide comprises a polypeptide of claim 4, or a polypeptide encoded by a nucleic acid of claim 1;

(b) providing a composition comprising a food or feed;

(c) expressing the nucleic acid to produce a recombinant alpha-amylase; and

(d) mixing the recombinant alpha-amylase and the feed-comprising composition, thereby producing a food or feed comprising a recombinant amylase.

2013201807 09 Sep 2015

21. A baking process:

(a) comprising use of a polypeptide of claim 4 or a polypeptide encoded by a nucleic acid of claim 1; or

(b) comprising use of a polypeptide of claim 4 or a polypeptide encoded by a nucleic acid of claim 1, wherein the baking process further comprises use of a second polypeptide having an amylase activity, an alpha amylase activity, or a beta amylase, or another enzyme.

22. A corn wet milling process:

(a) comprising use of a polypeptide of claim 4 or a polypeptide encoded by a nucleic acid of claim 1; or

(b) comprising use of a polypeptide of claim 4 or a polypeptide encoded by a nucleic acid of claim 1, wherein the corn wet milling process further comprises use of a second polypeptide having amylase activity, an alpha amylase activity, or a beta amylase, or another enzyme.

23. A dry milling process:

(a) comprising use of a polypeptide of claim 4 or a polypeptide encoded by the nucleic acid of claim 1; or

(b) comprising use of a polypeptide of claim 4 or a polypeptide encoded by the nucleic acid of claim 1, wherein the process further comprises use of a second polypeptide having amylase activity, an alpha amylase activity, or a beta amylase, or another enzyme.

24. A drilling process:

(a) comprising use of a polypeptide of claim 4 or a polypeptide encoded by a nucleic acid of claim 1; or

2013201807 09 Sep 2015

(b) comprising use of a polypeptide of claim 4 or a polypeptide encoded by a nucleic acid of claim 1, wherein the process further comprises use of a second polypeptide having amylase activity, an alpha amylase activity, or a beta amylase, or another enzyme.

25. A method for making an alcohol:

(a) comprising use of a polypeptide of claim 4 or a polypeptide encoded by a nucleic acid of claim 1; or

(b) comprising use of a polypeptide of claim 4 or a polypeptide encoded by a nucleic acid of claim 1, wherein the process further comprises use of a second polypeptide having amylase activity, an alpha amylase activity, or a beta amylase, or another enzyme.

26. An isolated, synthetic, or recombinant nucleic acid according to claim 1, substantially as hereinbefore described.

27. An isolated, synthetic, or recombinant polypeptide having alpha-amylase activity according to claim 4, substantially as hereinbefore described.

2013201807 18 Mar 2013

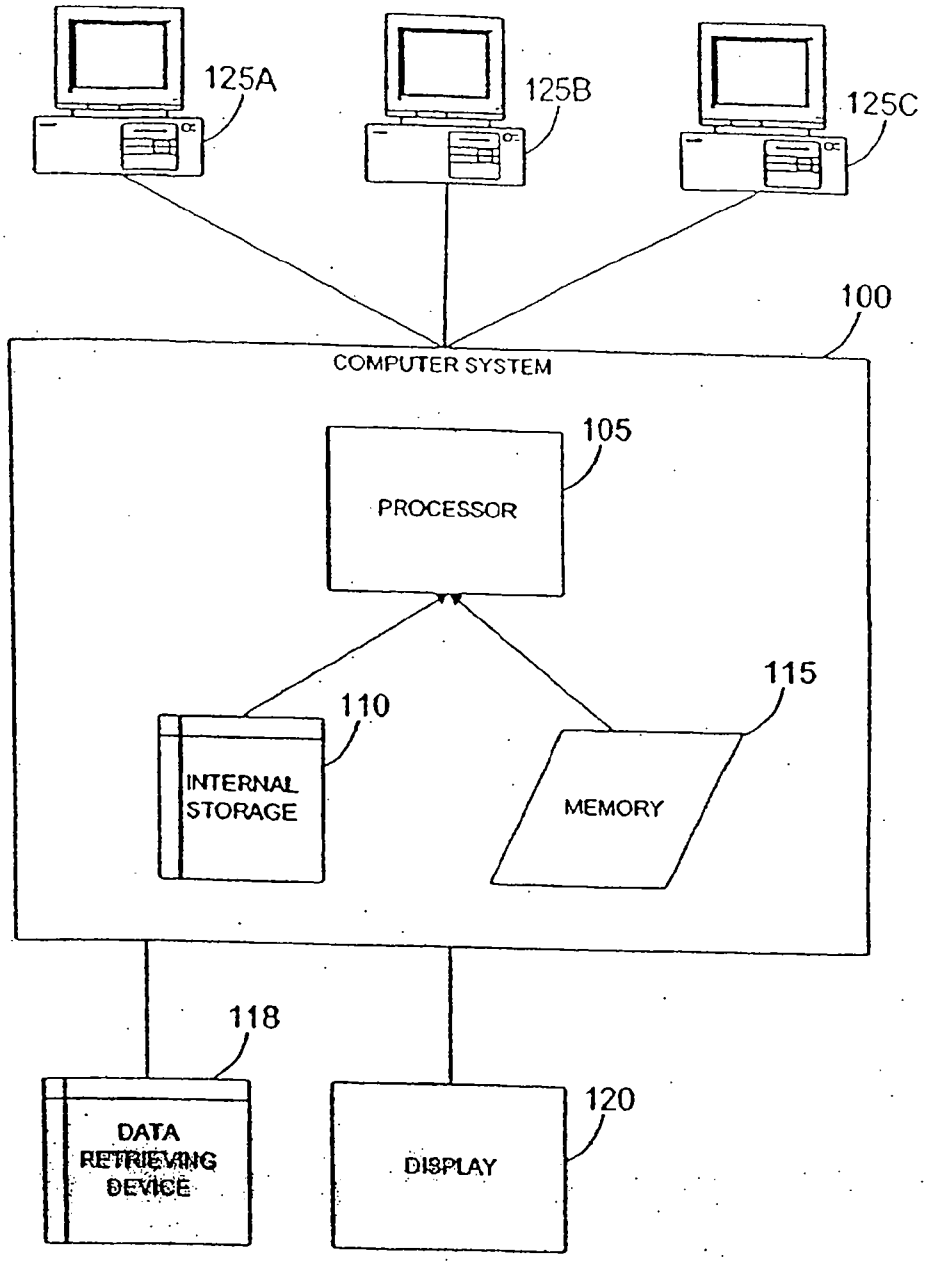


FIGURE 1

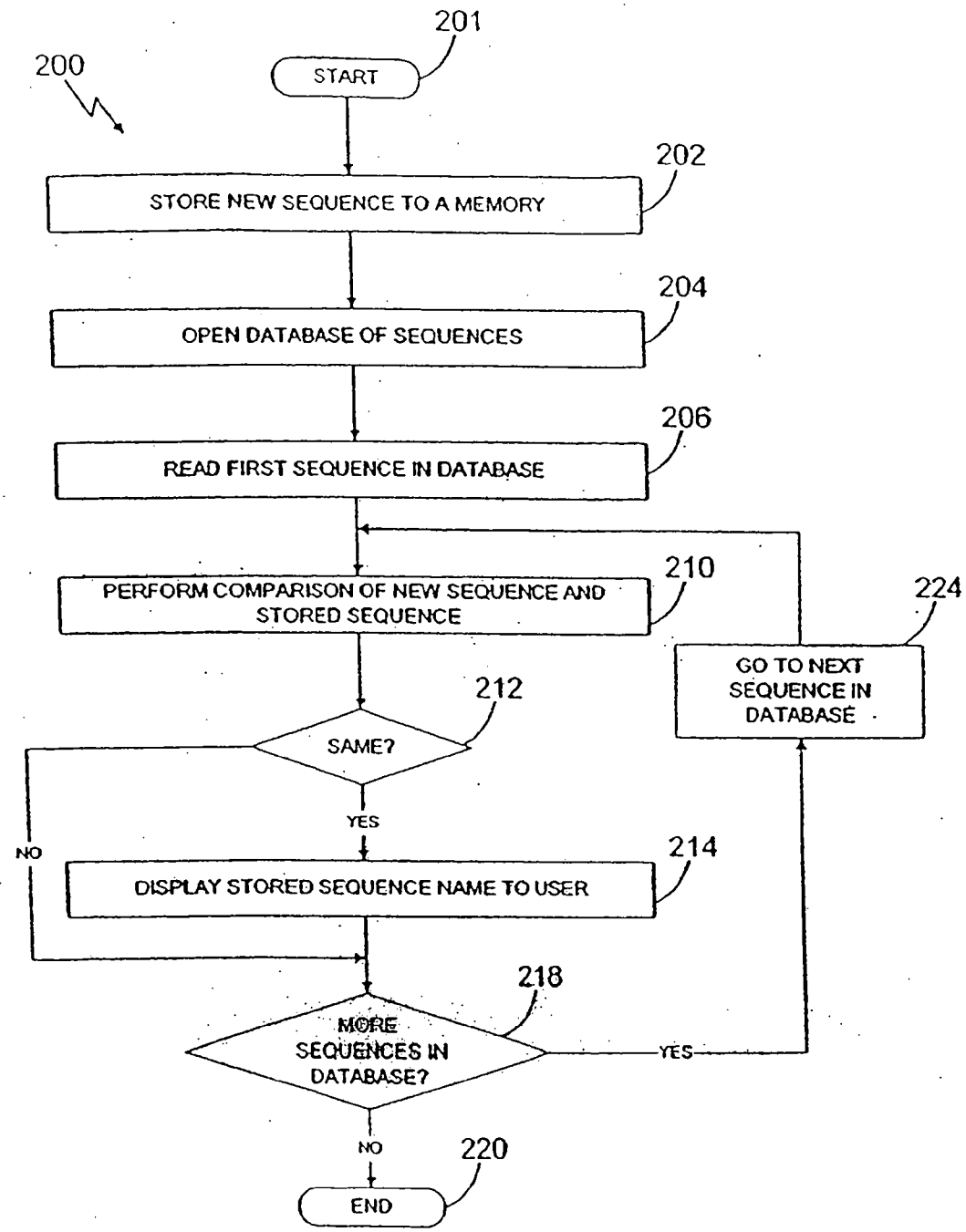

200
201
START
202
STORE NEW SEQUENCE TO A MEMORY
204
OPEN DATABASE OF SEQUENCES
206
READ FIRST SEQUENCE IN DATABASE
210
PERFORM COMPARISON OF NEW SEQUENCE AND STORED SEQUENCE
212
SAME?
YES
NO
214
DISPLAY STORED SEQUENCE NAME TO USER
218
MORE SEQUENCES IN DATABASE?
YES
NO
220
END
224
GO TO NEXT SEQUENCE IN DATABASE

FIGURE 2

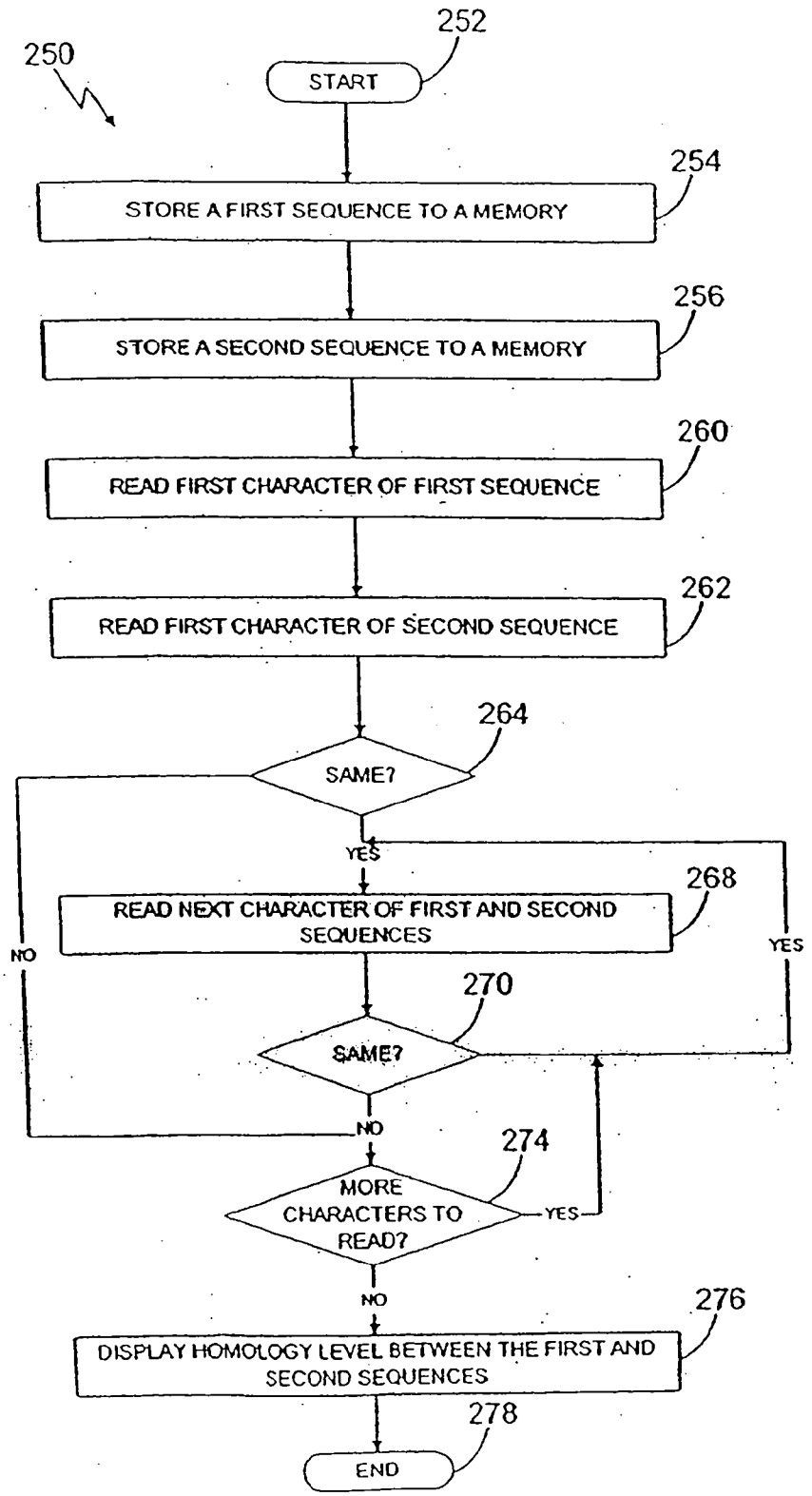

250
252
START
254
STORE A FIRST SEQUENCE TO A MEMORY
256
STORE A SECOND SEQUENCE TO A MEMORY
260
READ FIRST CHARACTER OF FIRST SEQUENCE
262
READ FIRST CHARACTER OF SECOND SEQUENCE
264
SAME?
NO
YES
268
READ NEXT CHARACTER OF FIRST AND SECOND SEQUENCES
YES
270
SAME?
NO
274
MORE CHARACTERS TO READ?
YES
NO
276
DISPLAY HOMOLOGY LEVEL BETWEEN THE FIRST AND SECOND SEQUENCES
278
END

FIGURE 3

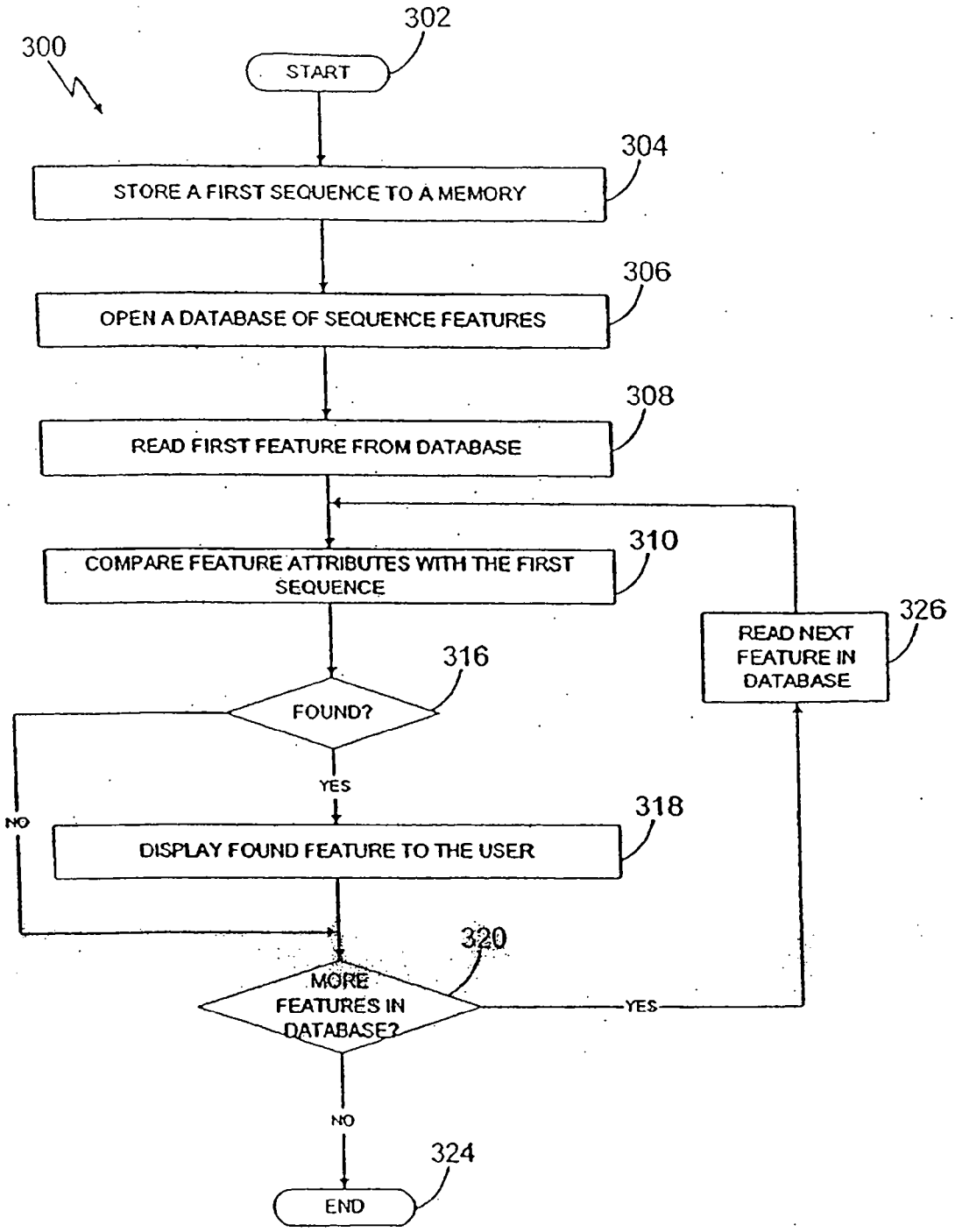

300
302
START
304
STORE A FIRST SEQUENCE TO A MEMORY
306
OPEN A DATABASE OF SEQUENCE FEATURES
308
READ FIRST FEATURE FROM DATABASE
310
COMPARE FEATURE ATTRIBUTES WITH THE FIRST SEQUENCE
316
FOUND?
YES
NO
318
DISPLAY FOUND FEATURE TO THE USER
320
MORE FEATURES IN DATABASE?
YES
NO
324
END
326
READ NEXT FEATURE IN DATABASE

FIGURE 4

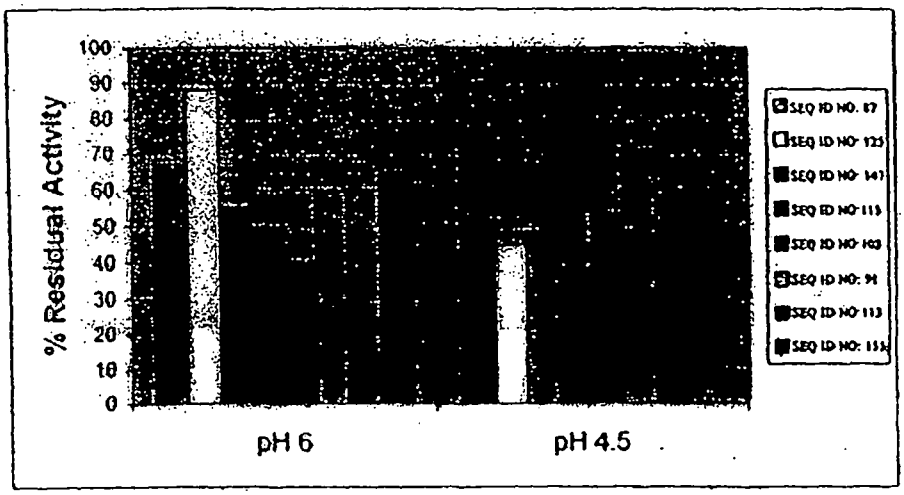

% Residual Activity
100
90
80
70
60
50
40
30
20
10
0
pH 6
pH 4.5
SEQ ID NO: 87
SEQ ID NO: 125
SEQ ID NO: 141
SEQ ID NO: 113
SEQ ID NO: 91
SEQ ID NO: 133

FIGURE 5

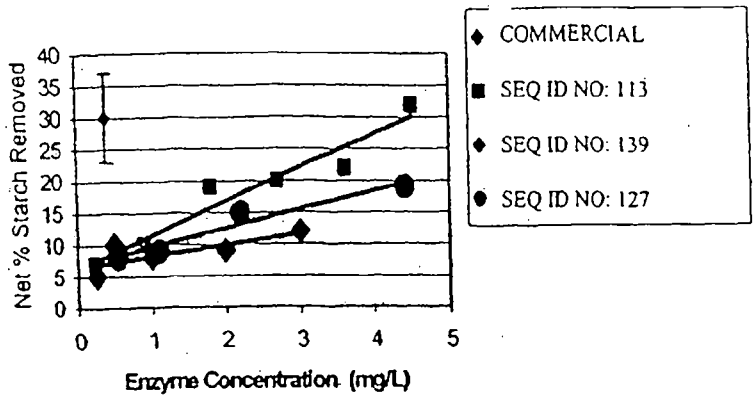

Net % Starch Removed
40
35
30
25
20
15
10
5
0
0
1
2
3
4
5
Enzyme Concentration (mg/L)
COMMERCIAL
SEQ ID NO: 113
SEQ ID NO: 139
SEQ ID NO: 127

FIGURE 6

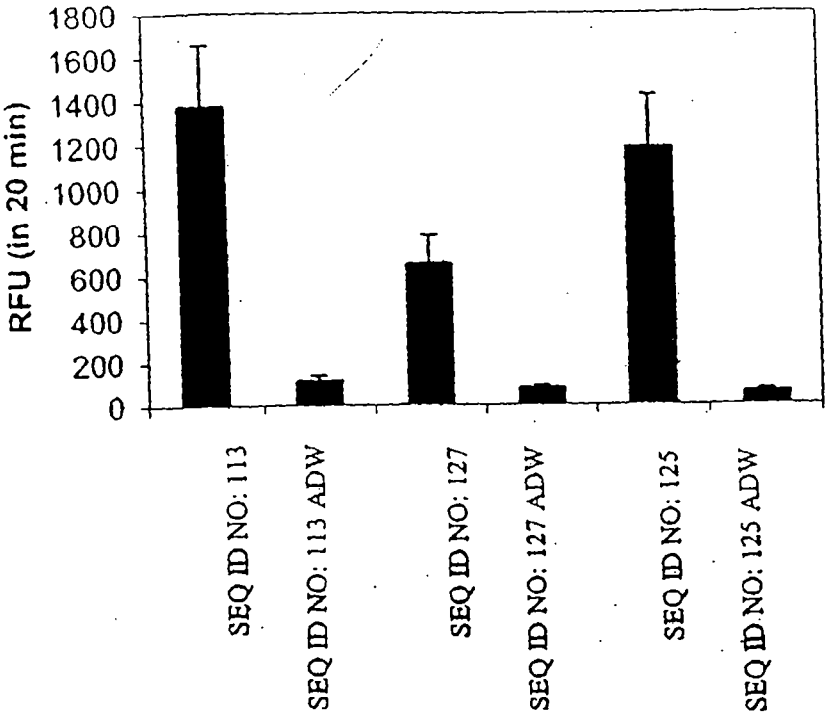

RFU (in 20 min)
1800
1600
1400
1200
1000
800
600
400
200
0
SEQ ID NO: 113
SEQ ID NO: 113 ADW
SEQ ID NO: 127
SEQ ID NO: 127 ADW
SEQ ID NO: 125
SEQ ID NO: 125 ADW

FIGURE 7

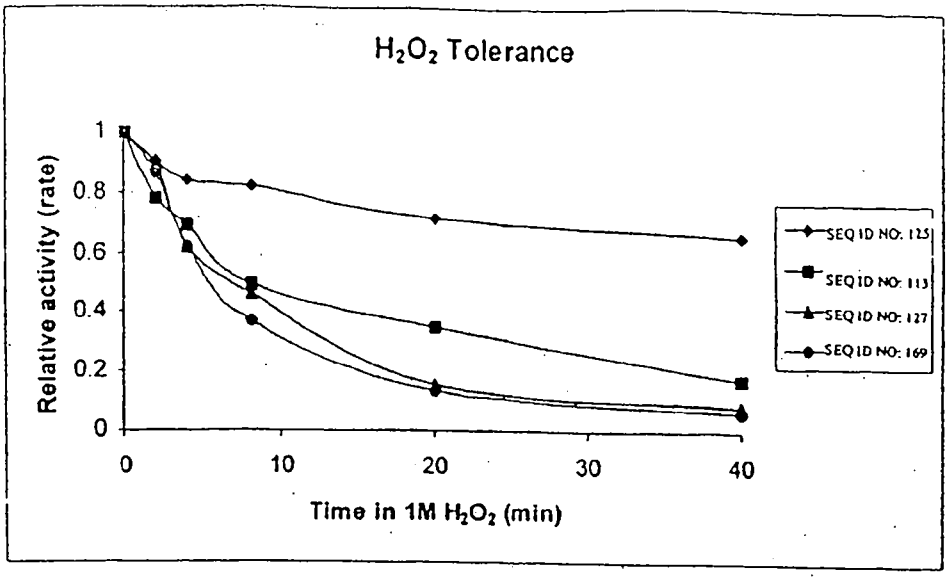

H_2O_2 Tolerance
Relative activity (rate)
1
0.8
0.6
0.4
0.2
0
0
10
20
30
40
Time in 1M H_2O_2 (min)
SEQ ID NO: 125
SEQ ID NO: 113
SEQ ID NO: 127
SEQ ID NO: 169

FIGURE 8

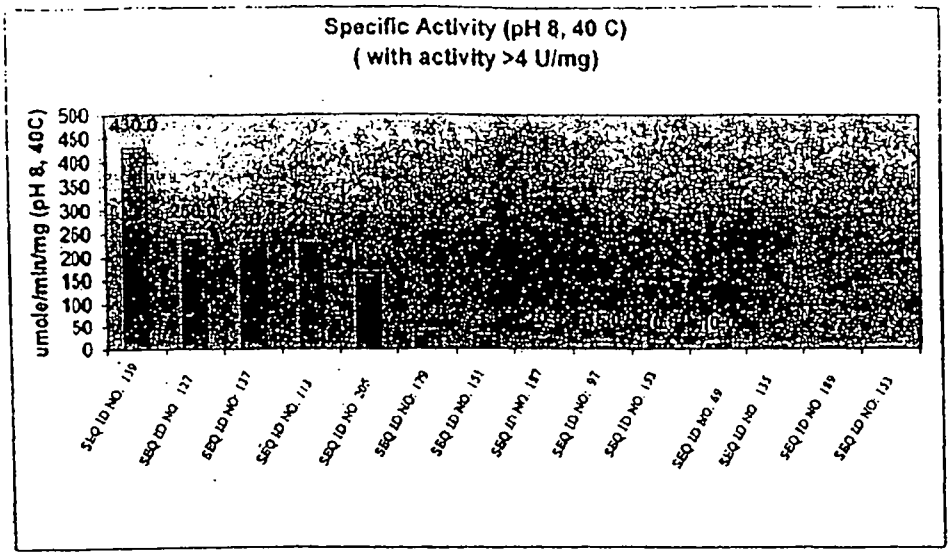


FIGURE 9A

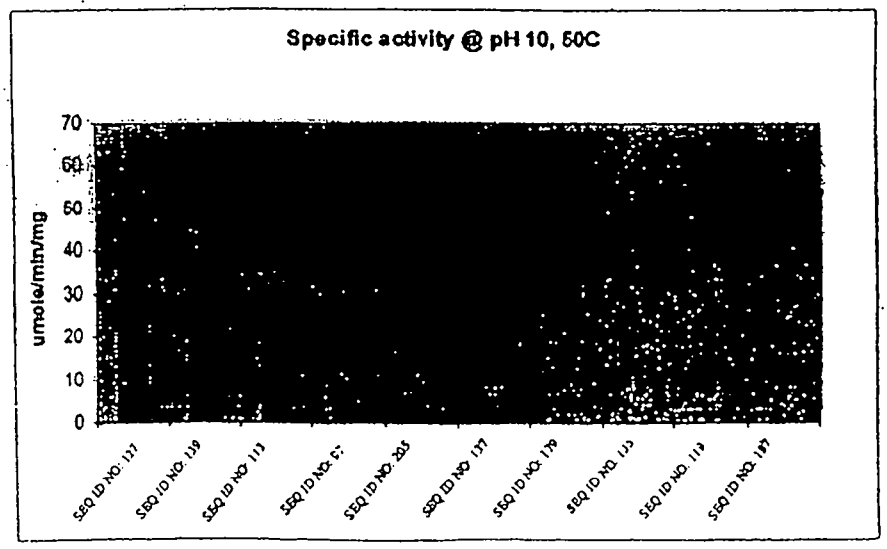


FIGURE 9B

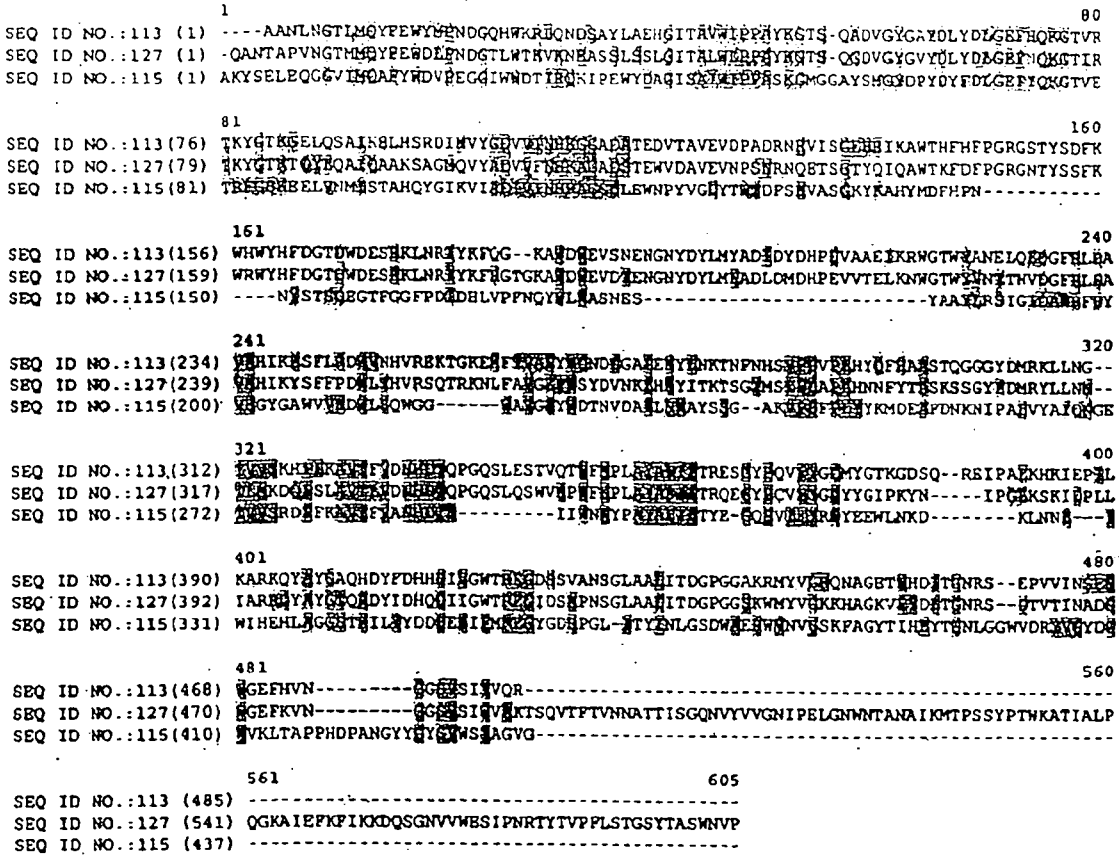


FIGURE 10

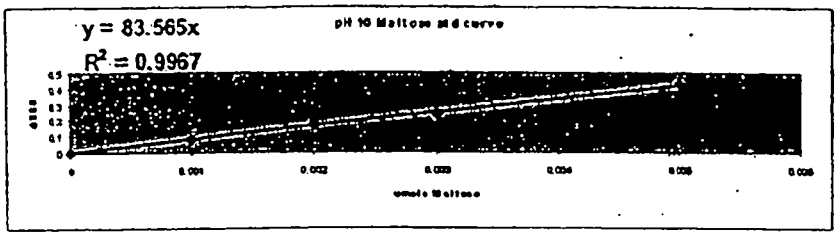


FIGURE 11

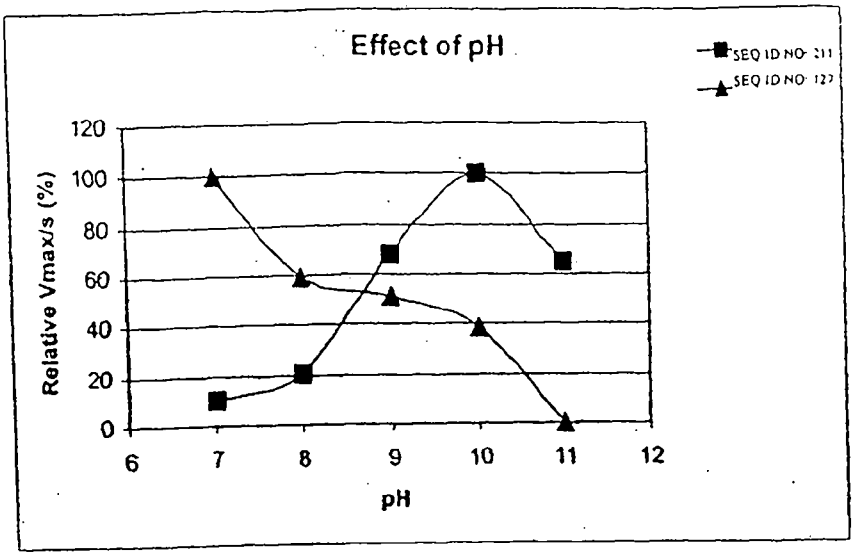


FIGURE 12

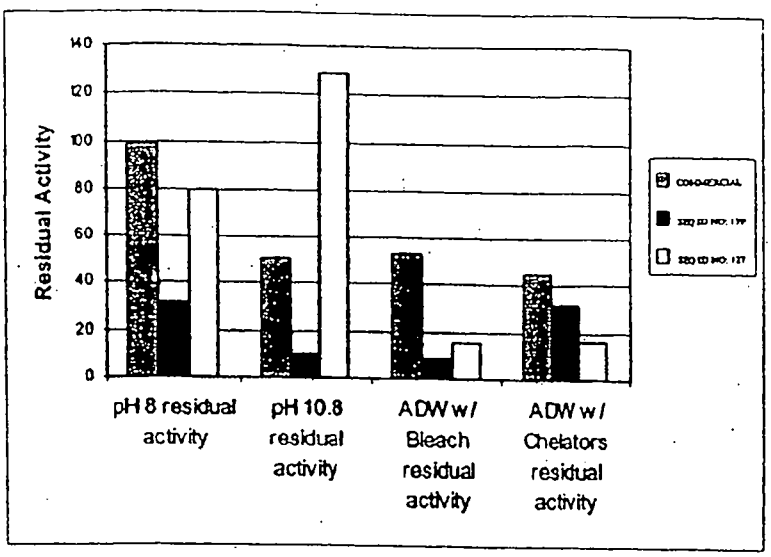

Residual Activity
140
120
100
80
60
40
20
0
pH 8 residual activity
pH 10.8 residual activity
ADW w/ Bleach residual activity
ADW w/ Chelators residual activity

FIGURE 13

```
                1                                                   50
SEQ ID NO: 81   ---------- -------MKK FVALFITMFF VVSMAVV... ..AQPASAAK
pyro            ---------- -------MKK FVALLITMFF VVSMAAV... ..AQPASAAK
pyro2           ---------- -----VNIKK LTPLLTLLLF FI...VL... ..ASPVSAAK
thermo          SESQCTATCT WRVVYMSAKK LLALLFVLAV LVGVAVIPAR VGIAPVSAGA
thermo2         ---------- ---------- --------MA RKVLVALLVF LVVLSVSAVP
Consensus       ---------- ---------- ---------- ---------- ------SA--

                51                                                  100
SEQ ID NO: 81   YS..ELEEGG VIMQAFYWDV PGGGIWWDTI RSKIPEWYEA GISAIWIPPA
pyro            YS..ELEEGG VIMQAFYWDV PAGGIWWDTI RSKIPEWYEA GISAIWIPPA
pyro2           YL..ELEEGG VIMQAFYWDV PGGGIWWDHI RSKIPEWYEA GISAIWLPPP
thermo          TSRPSLEEGG VIMQAFYWDV PAGGIWWDTI RSKIPDWASA GISAIWIPPA
thermo2         AKAETLENGG VIMQAFYWDV PGGGIWWDTI AQKIPDWASA GISAIWIPPA
Consensus       -----LE-GG VIMQAFYWDV P-GGIWWD-I --KIP-W--A GISAIW-PP-
                  Sense primer
                101                                                 150
SEQ ID NO: 81   SKGMSGGYSM GYDPYDFFDL GEYNQKGTIE TRFGSKQELI NMINTAHAYG
pyro            SKGMGGAYSM GYDPYDFFDL GEYNQKGTVE TRFGSKQELI NMINTAHAYG
pyro2           SKGMSGGYSM GYDPYDYFDL GEYYQKGTVE TRFGSKEELV RLIQTAHAYG
thermo          SKGMSGAYSM GYDPYDFFDL GEYYQKGTVE TRFGSKQELI NMINTAHSYG
thermo2         SKGMSGGYSM GYDPYDFFDL GEYYQKGSVE TRFGSKEELV NMINTAHAHN
Consensus       SKGM-G-YSM GYDPYD-FDL GEY-QKG--E TRFGSK-EL- --I-TAH---

                151                                                 200
SEQ ID NO: 81   IKVIADIVIN HRAGGDLEWN PFVGDYTWTD FSKVASGKYT ANYLDFHPNE
pyro            IKVIADIVIN HRAGGDLEWN PFVGDYTWTD FSKVASGKYT ANYLDFHPNE
pyro2           IKVIADVVIN HRAGGDLEWN PFVGDYTWTD FSKVASGKYT ANYLDFHPNE
thermo          IKVIADIVIN HRAGGDLEWN PFTNSYTWTD FSKVASGKYT ANYLDFHPNE
thermo2         MKVIADIVIN HRAGGDLEWN PFTNSYTWTD FSKVASGKYT ANYLDFHPNE
Consensus       -KVIAD-VIN HRAGGDLEWN PF---YTWTD FSKVASGKYT ANYLDFHPNE

                201                                                 250
SEQ ID NO: 81   VKCCDEGTFG GFPDIAHEKS WDQHWLWASD ESYAAYLRSI GVDAWRFDYV
pyro            VKCCDEGTFG GFPDIAHEKE WDQHWLWASD ESYAAYLRSI GVDAWRFDYV
pyro2           LHCCDEGTFG GFPDICHHKE WDQYWLWKSN ESYAAYLRSI GFDGWRFDYV
thermo          VKCCDEGTFG GFPDIAHEKS WDQYWLWASQ KSYAAYLRSI GIDAWRFDYV
thermo2         LHAGDSGTFG GYPDICHDKS WDQHWLWASN ESYAAYLRSI GIDAWRFDYV
Consensus       ----D-GTFG G-PDI-H-K- WDQ-WLW-S- -SYAAYLRSI G-D-WRFDYV

                251                                                 300
SEQ ID NO: 81   KGYGAWVVKD WLNWWGGWAV GEYWDTNVDA LLNWAYSSGA KVFDFPLYYK
pyro            KGYGAWVVKD WLNWWGGWAV GEYWDTNVDA LLNWAYSSGA KVFDFPLYYK
pyro2           KGYGAWVVRD WLNWWGGWAV GEYWDTNVDA LLSWAYESGA KVFDFPLYYK
thermo          KGYGAWVVKD WLKWW.ALAV GEYWDTNVDA LLNWAYSSGA KVFDFPLYYK
thermo2         KGYAPWVVKN WLNRWGGWAV GEYWDTNVDA LLSWAYDSGA KVFDFPLYYK
Consensus       KGY--WVV-- WL--W---AV GEYWDTNVDA LL-WAY-SGA KVFDFPLYYK

                301                                                 350
SEQ ID NO: 81   MDEAFDNKNI PALVSALQNG QTVVSRDPFK AVTFVANHDT DIIWNKYLAY
pyro            MDEAFDNTNI PALVDALQNG GTVVSRDPFK AVTFVANHDT DIIWNKYPAY
pyro2           MDEAFDNNNI PALVYALQNG QTVVSRDPFK AVTFVANHDT DIIWNKYPAY
thermo          MDEAFDNKNI PALVSALQNG QTVVSRDPFK AVTFVANHDT DIIWNKYPAY
thermo2         MDEAFDNNNI PALVDALKNG GTVVSRDPFK AVTFVANHDT NIIWNKYPAY
Consensus       MDEAFDN-NI PALV-AL-NG -TVVSRDPFK AVTFVANHDT -IIWNKY-AY
```

FIGURE 14A

```
                351                                                 400
SEQ ID NO: 81   AFILTYEGQP VIFYRDYEEW LNKDRLNNLI WIHDHLAGGS TSIVYYDSDE
      pyro      AFILTYEGQP VIFYRDYEEW LNKDKLNNLI WIHDHLAGGS TSIVYYDSDE
     pyro2      AFILTYEGQP VIFYRDFEEW LNKDKLINLI WIHDHLAGGS TTIVYYDNDE
     thermo     AFILTYEGQP VIFYRDYEEW LNKDRLKNLI WIHNNLAGGS TSIVYYDNDE
    thermo2     AFILTYEGQP AIFYRDYEEW LNKDRLRNLI WIHDHLAGGS TDIIYYDSDE
   Consensus    AFILTYEGQP -IFYRD-EEW LNKD-L-NLI WIH--LAGGS T-I-YYD-DE

                401                                                 450
SEQ ID NO: 81   MIFVRNGYGS KPGLITYINL GSSKVGRWVY VPKFAGACIH EYTGNLGGWV
      pyro      LIFVRNGDSK RPGLITYINL GSSKVGRWVY VPKFAGACIH EYTGNLGGWV
     pyro2      LIFVRNGDSR RPGLITYINL SPNWVGRWVY VPKFAGACIH EYTGNLGGWV
     thermo     LIFVRNGYGN KPGLITYINL GSSKVGRWVY VPKFAGSCIH EYTGNLGGWV
    thermo2     LIFVRNGYGD KPGLITYINL GSSKAGRWVY VPKFAGSCIH EYTGNLGGWI
   Consensus    -IFVRNG--- -PGLITYINL -----GRWVY VPKFAG-CIH EYTGNLGGW-

                451                                 486
SEQ ID NO: 81   DKYVYSSGWV YFEAPAYDPA NGQYGYSVWS YCGVG*
      pyro      DKYVESSGWV YLEAPAYDPA SGQYGYTVWS YCGVG*
     pyro2      DKRVDSSGWV YLEAPPHDPA NGYYGYSVWS YCGVG*
     thermo     DKYVGSNGWV YLEAPAHDPA KGQYGYSVWS YCGVG*
    thermo2     DKWVDSSGRV YLEAPAHDPA NGQYGYSVWS YCGVG*
   Consensus    DK-V-S-G-V Y-EAP--DPA -G-YGY-VWS YCGVG*
                                      Antisense primer
```

FIGURE 14A
(cont.)

```
                1                                                   50
SEQ ID NO: 81   ---------- ---------- ----MKKFVA LFITMFFVVS MAVVAQPASA
pyro            ---------- ---------- ----MKKFVA LLITMFFVVS MAAVAQPASA
SEQ ID NO: 73   ---------- ---------- ---------- ---------- ----------
thermo2         ---------- ---------- --------MA RKVLVALLVF LVVLSVSAVP
SEQ ID NO: 75   ---------- ---------- ---------- ---------- ----------
SEQ ID NO: 77   ---------- ---------- ---------- ---------- ----------
SEQ ID NO: 83   ---------- ---------- ---------- ---------- ----------
SEQ ID NO: 85   ---------- ---------- ---------- ---------- ----------
SEQ ID NO: 79   ---------- ---------- ----MKPAKL LVFVLVVSIL AGLYAQPAGA
thermo          SESQCTATCT WRVVYMSAKK LLALLFVLAV LVGVAVIPAR VGIAPVSAGA
pyro2           ---------- ---------- -----VNIKK LTPLLTLLLF FIVLASPVSA
CLONE A         ---------- ---------- ---MRRSARV LVLIIAFFLL AGIYYPSTSA
Consensus       ---------- ---------- ---------- ---------- ----------

                51                                                  100
SEQ ID NO: 81   AKYSELEEGG VIMQAFYWDV PGGGIWWDTI RSKIPEWYEA GISAIWIPPA
pyro            AKYSELEEGG VIMQAFYWDV PAGGIWWDTI RSKIPEWYEA GISAIWIPPA
SEQ ID NO: 73   ---MALEEGG LIMQAFYWDV PGGGIWWDTI AQKIPDWASA GISAIWIPPA
thermo2         AKAETLENGG VIMQAFYWDV PGGGIWWDTI AQKIPDWASA GISAIWIPPA
SEQ ID NO: 75   ---MALEEGG LIMQAFYWDV PMGGIWWDTI AQKIPDWASA GISAIWIPPA
SEQ ID NO: 77   ---MALEEGG LIMQAFYWDV PMGGIWWDTI AQKIPDWASA GISAIWIPPA
SEQ ID NO: 83   ---MALEEGG LIMQAFYWDV PGGGIWWDTI AQKIPEWASA GISAIWIPPA
SEQ ID NO: 85   ---MALEEGG LIMQAFYWDV PGGGIWWDTI AQKIPEWASA GISAIWIPPA
SEQ ID NO: 79   AKYLELEEGG VIMQAFYWDV PSGGIWWDTI RQKIPEWYDA GISAIWIPPA
thermo          TSRPSLEEGG VIMQAFYWDV PAGGIWWDTI RSKIPDWASA GISAIWIPPA
pyro2           AKYLELEQGG VIMQAFYWDV PEGGIWWDTI RQKIPEWYDA GISAIWIPPA
CLONE A         AKYSELEQGG VIMQAFYWDV P-GGIWWD-I --KIP-W--A GISAIW-PP-
Consensus       --------GG -IMQAFYWDV P-GGIWWD-I --KIP-W--A GISAIW-PP-

                101                                                 150
SEQ ID NO: 81   SKGMSGGYSM GYDPYDFFDL GEYNQKGTIE TRFGSKQELI NMINTAHAYG
pyro            SKGMGGAYSM GYDPYDFFDL GEYNQKGTVE TRFGSKQELI NMINTAHAYG
SEQ ID NO: 73   SKGMSGGYSM GYDPYDFFDL GEYYQKGSVE TRFGSKEELV NMINTAHAHN
thermo2         SKGMSGGYSM GYDPYDFFDL GEYYQKGSVE TRFGSKEELV NMINTAHAHN
SEQ ID NO: 75   SKGMSGGYSM GYDPYDYFDL GEYYQKGTVE TRFGSKQELI NMINTAHAYG
SEQ ID NO: 77   SKGMSGGYSM GYDPYDYFDL GEYYQKGTVE TRFGSKQELI NMINTAHAYG
SEQ ID NO: 83   SKGMSGGYSM GYDPYDFFDL GEYYQKGTVE TRFGSKEELV NMINTAHSYG
SEQ ID NO: 85   SKGMSGGYSM GYDPYDFFDL GEYYQKGTVE TRFGSKEELV NMINTAHSYG
SEQ ID NO: 79   SKGMGGAYSM GYDPYDFFDL GEYDQKGTVE TRFGSKQELV NMINTAHAYG
thermo          SKGMSGAYSM GYDPYDFFDL GEYYQKGTVE TRFGSKQELI NMINTAHSYG
pyro2           SKGMSGGYSM GYDPYDYFDL GEYYQKGTVE TRFGSKEELV RLIQTAHAYG
CLONE A         SKGMGGAYSM GYDPYDYFDL GEFYQKGTVE TRFGSKEELV NMISTAHQYG
Consensus       SKGM-G-YSM GYDPYD-FDL GE--QKG--E TRFGSK-EL- --I-TAH---

                151                                                 200
SEQ ID NO: 81   IKVIADIVIN HRAGGDLEWN PFVGDYTWTD FSKVASGKYT ANYLDFHPNE
pyro            IKVIADIVIN HRAGGDLEWN PFVGDYTWTD FSKVASGKYT ANYLDFHPNE
SEQ ID NO: 73   MKVIADIVIN HRAGGDLEWN PFTNSYTWTD FSKVASGKYT ANYLDFHPNE
thermo2         MKVIADIVIN HRAGGDLEWN PFTNSYTWTD FSKVASGKYT ANYLDFHPNE
SEQ ID NO: 75   MKVIADIVIN HRAGGDLEWN PFVNDYTWTD FSKVASGKYT ANYLDFHPNE
SEQ ID NO: 77   MKVIADIVIN HRAGGDLEWN PFVNDYTWTD FSKVASGKYT ANYLDFHPNE
SEQ ID NO: 83   IKVIADIVIN HRAGGDLEWN PFVNDYTWTD FSKVASGKYT ANYLDFHPNE
SEQ ID NO: 85   IKVIADIVIN HRAGGGLEWN PFVNDYTWTD FSKVASGKYT ANYLDFHPNE
SEQ ID NO: 79   IKVIADIVIN HRAGGDLEWN PFVNDYTWTD FSKVASGKYT ANYLDFHPNE
thermo          IKVIADIVIN HRAGGDLEWN PFTNSYTWTD FSKVASGKYT ANYLDFHPNE
pyro2           IKVIADVVIN HRAGGDLEWN PFVGDYTWTD FSKVASGKYT ANYLDFHPNE
CLONE A         IKVIADIVIN HRAGGDLEWN PYVGDYTWTD FSKVASGKYK AHYMDFHPNN
Consensus       -KVIAD-VIN HRAGG-LEWN P----YTWTD FSKVASGKY- A-Y-DFHPN-
```

FIGURE 14B

```
                201                                                                250
SEQ ID NO: 81   VKCCDEGTFG GFPDIAHEKS WDQHWLWASD ESYAAYLRSI GVDAWRFDYV
          pyro  VKCCDEGTFG GFPDIAHEKE WDQHWLWASD ESYAAYLRSI GVDAWRFDYV
SEQ ID NO: 73   LHAGDSGTFG GYPDICHDKS WDQHWLWASN ESYAAYLRSI GIDAWRFDYV
       thermo2  LHAGDSGTFG GYPDICHDKS WDQHWLWASN ESYAAYLRSI GIDAWRFDYV
SEQ ID NO: 75   LHAGDSGTFG GYPDICHDKS WDQYWLWASQ ESYAAYLRSI GIDAWRFDYV
SEQ ID NO: 77   LHAGDSGTFG GYPDICHDKS WDQYWLWASQ ESYAAYLRSI GIDAWRFDYV
SEQ ID NO: 83   LHCCDEGTFG GYPDICHDKS WDQYWLWASS ESYAAYLRSI GVDAWRFDYV
SEQ ID NO: 85   LHCCDEGTFG GYPDICHDKS WDQYWLWASS ESYAAYLRSI GVDAWCFDYV
SEQ ID NO: 79   VKCCDEGTFG GFPDIAHEKS WDQYWLWASN ESYAAYLRSI GVDAWRFDYV
       thermo   VKCCDEGTFG GFPDIAHEKS WDQYWLWASQ KSYAAYLRSI GIDAWRFDYV
        pyro2   LHCCDEGTFG GFPDICHHKE WDQYWLWKSN ESYAAYLRSI GFDGWRFDYV
       CLONE A  YSTSDEGTFG GFPDIDHLVP FNQYWLWASN ESYAAYLRSI GIDAWRFDYV
    Consensus   ----D-GTFG G-PDI-H---  --Q-WLW-S-  -SYAAYLRSI G-D-W-FDYV

                251                                                                300
SEQ ID NO: 81   KGYGAWVVKD WLNWWGGWAV GEYWDTNVDA LLNWAYSSGA KVFDFPLYYK
          pyro  KGYGAWVVKD WLNWWGGWAV GEYWDTNVDA LLNWAYSSGA KVFDFPLYYK
SEQ ID NO: 73   KGYAPWVVKN WLNRWGGWAV GEYWDTNVDA LLSWAYDSGA KVFDFPLYYK
       thermo2  KGYAPWVVKN WLNRWGGWAV GEYWDTNVDA LLSWAYDSGA KVFDFPLYYK
SEQ ID NO: 75   KGYAPWVVRD WLNWWGGWAV GEYWDTNVDA VLNWAYSSGA KVFDFALYYK
SEQ ID NO: 77   KGYAPWVVKD WLNWWGGWAV GEYWDTNVDA VLNWAYSSGA KVFDFALYYK
SEQ ID NO: 83   KGYGAWVVND WLSWWGGWAV GEYWDTNVDA LLNWAYSSGA KVFDFPLYYK
SEQ ID NO: 85   KGYGAWVVND WLSWWGGWAV GEYWDTNVDA LLNWAYNSGA KVFDFPLYYK
SEQ ID NO: 79   KGYGAWVVKD WLDWWGGWAV GEYWDTNVDA LLNWAYSSDA KVFDFPLYYK
       thermo   KGYGAWVVKD WLKWW.ALAV GEYWDTNVDA LLNWAYSSGA KVFDFPLYYK
        pyro2   KGYGAWVVRD WLNWWGGWAV GEYWDTNVDA LLSWAYESGA KVFDFPLYYK
       CLONE A  KGYGAWVVKD WLSQWGGWAV GEYWDTNVDA LLNWAYSSGA KVFDFPLYYK
    Consensus   KGY--WVV--  WL--W---AV GEYWDTNVDA -L-WAY-S-A KVFDF-LYYK

                301                                                                350
SEQ ID NO: 81   MDEAFDNKNI PALVSALQNG QTVVSRDPFK AVTFVANHDT DIIWNKYLAY
          pyro  MDEAFDNTNI PALVDALQNG GTVVSRDPFK AVTFVANHDT DIIWNKYPAY
SEQ ID NO: 73   MDEAFDNNNI PALVDALKNG GTVVSRDPFK AVTFVANHDT NIIWNKYPAY
       thermo2  MDEAFDNNNI PALVDALKNG GTVVSRDPFK AVTFVANHDT NIIWNKYPAY
SEQ ID NO: 75   MDEAFDNNNI PALVDALRYG QTVVSRDPFK AVTFVANHDT DIIWNKYPAY
SEQ ID NO: 77   MDEAFDNNNI PALVDALRYG QTVVSRDPFK AVTFVANHDT DIIWNKYPAY
SEQ ID NO: 83   MDEAFDNTNI PALVDALRYG QTVVSRDPFK AVTFVANHDT DIIWNKYPAY
SEQ ID NO: 85   MDEAFDNTNI PALVYALKNG GTVVSRDPFK AVTFVANHDT DIIWNKYPAY
SEQ ID NO: 79   MDAAFDNKNI PALVEALKNG GTVVSRDPFK AVTFVANHDT DIIWNKYPAY
       thermo   MDEAFDNKNI PALVSALQNG QTVVSRDPFK AVTFVANHDT DIIWNKYPAY
        pyro2   MDEAFDNNNI PALVYALQNG QTVVSRDPFK AVTFVANHDT DIIWNKYPAY
       CLONE A  MDEAFDNKNI PALVYALQNG GTVVSRDPFK AVTFVANHDT NIIWNKYPAY
    Consensus   MD-AFDN-NI PALV-A---G -TVVSRDPFK AVTFVANHDT -IIWNKY-AY

                351                                                                400
SEQ ID NO: 81   AFILTYEGQP VIFYRDYEEW LNKDRLNNLI WIHDHLAGGS TSIVYYDSDE
          pyro  AFILTYEGQP VIFYRDYEEW LNKDKLNNLI WIHDHLAGGS TSIVYYDSDE
SEQ ID NO: 73   AFILTYEGQP AIFYRDYEEW LNKDRLRNLI WIHDHLAGGS TDIIYYDSDE
       thermo2  AFILTYEGQP AIFYRDYEEW LNKDRLRNLI WIHDHLAGGS TDIIYYDSDE
SEQ ID NO: 75   AFILTYEGQP TIFYRDYEEW LNKDKLKNLI WIHDNLAGGS TDIVYYDNDE
SEQ ID NO: 77   AFILTYEGQP TIFYRDYEEW LNKDKLKNLI WIHDNLAGGS TDIVYYDNDE
SEQ ID NO: 83   AFILTYEGQP VIFYRDYEEW LNKDKLNNLI WIHDHLAGGS TDIVYYDSDE
SEQ ID NO: 85   AFILTYEGQP VIFYRDYEEW LNKDKLNNLI WIHDHLAGGS TDIVYYDSDE
SEQ ID NO: 79   AFILTYEGQP TIFYRDYEEW LNKDRLKNLI WIHDHLAGGS TDIVYYDNDE
       thermo   AFILTYEGQP VIFYRDYEEW LNKDRLKNLI WIHNNLAGGS TSIVYYDNDE
        pyro2   AFILTYEGQP VIFYRDFEEW LNKDKLINLI WIHDHLAGGS TTIVYYDNDE
       CLONE A  AFILTYEGQP VIFYRDYEEW LNKDKLNNLI WIHEHLAGGS TKILYYDDDE
    Consensus   AFILTYEGQP -IFYRD-EEW LNKD-L-NLI WIH--LAGGS T-I-YYD-DE
```

FIGURE 14B

(cont.)

```
                401                                                 450
SEQ ID NO: 81   MIFVRNGYGS KPGLITYINL GSSKVGRWVY V.PKFAGACI HEYTGNLGGW
      pyro      LIFVRNGDSK RPGLITYINL GSSKVGRWVY V.PKFAGACI HEYTGNLGGW
SEQ ID NO: 73   LIFVRNGYGD KPGLITYINL GSSKAGRWVY V.PKFAGSCI HEYTGNLGGW
    thermo2     LIFVRNGYGD KPGLITYINL GSSKAGRWVY V.PKFAGSCI HEYTGNLGGW
SEQ ID NO: 75   LIFVRNGYGS KPGLITYINL GSSKAGRWVY V.PKFAGSCI HEYTGNLGGW
SEQ ID NO: 77   LIFVRNGYGS KPGLITYINL ASSKAGRWVY V.PKFAGSCI HEYTGNLGGW
SEQ ID NO: 83   LIFVRNGYGT KPGLITYINL GSSKVGRWVY V.PKFAGSCI HEYTGNLGGW
SEQ ID NO: 85   LIFVRNGYGT KPGLITYINL GSSKAGRWVY V.PKFAGSCI HEYTGSLGGW
SEQ ID NO: 79   LIFVRNGYGD KPGLITYINL GSSKAGRWVY V.PKFAGACI HEYTGNLGGW
      thermo    LIFVRNGYGN KPGLITYINL GSSKVGRWVY V.PKFAGSCI HEYTGNLGGW
      pyro2     LIFVRNGDSR RPGLITYINL SPNWVGRWVY V.PKFAGACI HEYTGNLGGW
      CLONE A   LIFMREGYGD RPGLITYINL GSDWABRWVN VGSKFAGYTI HEYTGNLGGW
    Consensus   -IF-R-G--- -PGLITYINL ------RWV- V--KFAG--I HEYTG-LGGW

                451                                     487
SEQ ID NO: 81   VDKYVYSSGW VYFEAPAYDP ANGQYGYSVW SYCGVG*
      pyro      VDKYVESSGW VYLEAPAYDP ASGQYGYTVW SYCGVG*
SEQ ID NO: 73   IDKWVDSSGR VYLEAPAHDP ANGQYGYSVW SYCGVG*
    thermo2     IDKWVDSSGR VYLEAPAHDP ANGQYGYSVW SYCGVG*
SEQ ID NO: 75   VDKWVDSSGW VYLEAPAHDP ANGQYGYSVW SYCGVG*
SEQ ID NO: 77   VDKWVDSSGW VYLEAPAHDP ANGQYGYSVW SYCGVG*
SEQ ID NO: 83   IDKYVSSSGW VYLEAPAHDP ANGYYGYSVW SYCGVG*
SEQ ID NO: 85   IDKYVSSSGW VYLEAPAHDP ANGQYGYSVW SYCGVG*
SEQ ID NO: 79   VDKWVDSSGW VYLEAPAHDP ANGYYGYSVW SYCGVG*
      thermo    VDKYVGSNGW VYLEAPAHDP AKGQYGYSVW SYCGVG*
      pyro2     VDKRVDSSGW VYLEAPPHDP ANGYYGYSVW SYCGVG*
      CLONE A   VDRYVQYDGW VKLTAPPHDP ANGYYGYSVW SYAGVG*
    Consensus   -D--V---G- V---AP--DP A-G-YGY-VW SY-GVG*
```

FIGURE 14B
(cont.)

```
                1                                                   50
SEQ ID NO: 83   ---------- ---------- ---------- ---------- ----------
SEQ ID NO: 85   ---------- ---------- ---------- ---------- ----------
SEQ ID NO: 75   ---------- ---------- ---------- ---------- ----------
SEQ ID NO: 77   ---------- ---------- ---------- ---------- ----------
SEQ ID NO: 73   ---------- ---------- ---------- ---------- ----------
SEQ ID NO: 79   ---ATGAAGC CTGCGAAACT CCTCGTCTTT GTGCTCGTAG TCTCTATCCT
SEQ ID NO: 81   ---ATGAAGA AGTTTGTCGC CCTGTTCATA ACCATGTTTT TCGTAGTGAG
     CLONE A    ATGAGGAGAT CCGCAAGGGT TTTGGTTCTG ATTATAGCGT TTTTCCTCCT
Consensus       ---------- ---------- ---------- ---------- ----------

                51                                                  100
SEQ ID NO: 83   ---------- ---------- ---------- ---------- ATGGCTCTGG
SEQ ID NO: 85   ---------- ---------- ---------- ---------- ATGGCTCTGG
SEQ ID NO: 75   ---------- ---------- ---------- ---------- ATGGCTCTGG
SEQ ID NO: 77   ---------- ---------- ---------- ---------- ATGGCTCTGG
SEQ ID NO: 73   ---------- ---------- ---------- ---------- ATGGCTCTGG
SEQ ID NO: 79   CGCGGGGCTC TACGCCCAGC CCGCGGGGGC GGCCAAGTAC CTGGAGCTCG
SEQ ID NO: 81   CATGGCAGTC GTTGCACAGC CAGCTAGCGC CGCAAAGTAT TCCGAGCTCG
     CLONE A    GGCGGGGATT TACTACCCCT CCACGAGTGC CGCGAAGTAC TCCGAGCTGG
Consensus       ---------- ---------- ---------- ---------- ----------

                101                                                 150
SEQ ID NO: 83   AAGAGGGCGG GCTCATAATG CAGGCCTTCT ACTGGGATGT TCCTGGAGGA
SEQ ID NO: 85   AAGAGGGCGG GCTTATAATG CAGGCATTCT ATTGGGACGT CCCAGGTGGA
SEQ ID NO: 75   AAGAGGGCGG GCTTATAATG CAGGCATTCT ACTGGGACGT CCCCATGGGA
SEQ ID NO: 77   AAGAGGGCGG GCTCATAATG CAGGCCTTCT ACTGGGACGT CCCCATGGGA
SEQ ID NO: 73   TAGAGGGCGG GCTTATAATG CAGGCCTTCT ACTGGGACGT CCCAGGTGGA
SEQ ID NO: 79   AAGAGGGCGG CGTCATAATG CAGGCGTTCT ACTGGGACGT GCCTTCAGGA
SEQ ID NO: 81   AAGAAGGCGG CGTTATAATG CAGGCCTTCT ACTGGGACGT CCCAGGTGGA
     CLONE A    AGCAGGGCGG AGTCATAATG CAGGCCTTCT ACTGGGACGT TCCGGAGGGA
Consensus       -----GGCGG --T-ATAATG CAGGC-TTCT A-TGGGA-GT -CC----GGA

                151                                                 200
SEQ ID NO: 83   GGAATCTGGT GGGACACAAT AGCTCAAAAG ATACCCGAAT GGGCAAGTGC
SEQ ID NO: 85   GGAATCTGGT GGGACACCAT AGCCCAGAAG ATACCCGAAT GGGCAAGTGC
SEQ ID NO: 75   GGAATCTGGT GGGACACGAT AGCCCAGAAG ATACCCGACT GGGCAAGCGC
SEQ ID NO: 77   GGAATCTGGT GGGACACGAT AGCCCAGAAG ATACCCGACT GGGCAAGCGC
SEQ ID NO: 73   GGAATCTGGT GGGACACCAT AGCCCAGAAG ATACCCGACT GGGCGAGCGC
SEQ ID NO: 79   GGAATATGGT GGGACACAAT ACGGCAGAAG ATACCGGAGT GGTACGATGC
SEQ ID NO: 81   GGAATCTGGT GGGACACCAT CAGGAGCAAG ATACCGGAGT GGTACGAGGC
     CLONE A    GGAATCTGGT GGGACACAAT ACGGCAGAAG ATCCCTGAAT GGTACGATGC
Consensus       GGAAT-TGGT GGGACAC-AT -------AAG AT-CC-GA-T GG-----GC

                201                                                 250
SEQ ID NO: 83   AGGAATCTCA GCGATATGGA TTCCACCAGC GAGTAAGGGC ATGAGCGGTG
SEQ ID NO: 85   AGGAATCTCA GCGATATGGA TTCCACCAGC GAGTAAGGGA ATGAGCGGTG
SEQ ID NO: 75   CGGGATTTCG GCGATATGGA TTCCCCCCGC GAGCAAGGGT ATGAGCGGCG
SEQ ID NO: 77   CGGGATTTCG GCGATATGGA TCCCTCCCGC GAGCAAGGGT ATGAGCGGCG
SEQ ID NO: 73   CGGGATTTCG GCAATATGGA TTCCTCCCGC GAGTAAGGGC ATGAGCGGCG
SEQ ID NO: 79   CGGAATCTCC GCAATATGGA TTCCCCCGGC GAGCAAGGGC ATGGGCGGCG
SEQ ID NO: 81   GGGAATATCC GCCATTTGGA TTCCGCCAGC CAGCAAGGGG ATGAGCGGCG
     CLONE A    AGGCATATCC GCCATCTGGA TACCCCCGGC GAGCAAGGGC ATGGGCGGGG
Consensus       -GG-AT-TC- GC-AT-TGGA T-CC-CC-GC -AG-AAGGG- ATG-GCGG-G
```

FIGURE 14C

```
                251                                                       300
SEQ ID NO: 83   GTTATTC CAT GGGCTACG AT CCCTACGAT T TCTTTGACCT CGGCGAGTAC
SEQ ID NO: 85   GTTATTC CAT GGGCTACG AT CCCTACGAT T TCTTTGACCT CGGCGAGTAC
SEQ ID NO: 75   GCTATTC GAT GGGCTACG AC CCCTACGAT T ATTTTGACCT CGGTGAGTAC
SEQ ID NO: 77   GCTATTC GAT GGGCTACG AC CCCTACGAT T ATTTTGACCT CGGTGAGTAC
SEQ ID NO: 73   GCTATTC GAT GGGCTACG AC CCCTACGAT T TCTTCGACCT CGGTGAGTAC
SEQ ID NO: 79   CCTATTC GAT GGGCTACG AC CCCTACGAC T TCTTTGACCT CGGTGAGTAC
SEQ ID NO: 81   GTTACTC GAT GGGCTACG AT CCCTACGAT T TCTTTGACCT CGGCGAGTAC
     CLONE A    CCTACTC GAT GGGCTACG AC CCCTACGAT T ACTTCGATCT GGGCGAGTTT
Consensus       --TA-TC -AT GGGCTACG A- CCCTACGA-T --TT-GA-CT -GG-GAGT--

                301                                                       350
SEQ ID NO: 83   TATCAGA AGG GGACAGTT GA GACGCGCTT C GGCTCAAAGG AAGAACTGGT
SEQ ID NO: 85   TATCAGA AGG GGACAGTT GA GACGCGCTT C GGCTCAAAGG AAGAACTGGT
SEQ ID NO: 75   TACCAGA AGG GAACGGTG GA AACAAGATT C GGCTCAAAGC AGGAGCTCAT
SEQ ID NO: 77   TACCAGA AGG GAACGGTG GA AACGAGGTT C GGCTCAAAGC AGGAGCTCAT
SEQ ID NO: 73   TACCAGA AGG GAAGCGTT GA GACCCGCTT C GGATCAAAAG AGGAGCTTGT
SEQ ID NO: 79   GACCAGA AGG GAACGGTA GA GACGCGCTT T GGCTCCAAGC AGGAGCTCGT
SEQ ID NO: 81   AACCAGA AGG GAACCATC GA AACGCGCTT T GGCTCTAAAC AGGAGCTCAT
     CLONE A    TACCAGA AGG GAACCGTT GA GACCCGCTT C GGCTCCAAGG AAGAGCTCGT
Consensus       -A-CAGA AGG G-A---T-GA --AC--G-TT- GG-TC-AA-- A-GA-CT--T

                351                                                       400
SEQ ID NO: 83   GAACATG ATA AACACCGC AC ACTCCTACG G CATAAAGGTG ATAGCAGACA
SEQ ID NO: 85   GAACATG ATA AACACCGC AC ACTCCTACG G CATAAAGGTG ATAGCGGACA
SEQ ID NO: 75   AAACATG ATA AACACCGC CC ACGCCTATG G CATGAAGGTA ATAGCCGATA
SEQ ID NO: 77   AAACATG ATA AACACCGC CC ACGCCTATG G CATGAAGGTA ATAGCCGATA
SEQ ID NO: 73   GAACATG ATA AACACCGC CC ATGCTCACA A CATGAAGGTC ATAGCGGACA
SEQ ID NO: 79   GAACATG ATA AACACCGC CC ACGCCTACG G CATCAAGGTC ATCGCAGACA
SEQ ID NO: 81   CAATATG ATA AACACGGC CC ATGCCTACG G CATAAAGGTC ATAGCGGACA
     CLONE A    CAACATG ATC TCCACGGC CC ACCAGTACG G CATCAAGGTT ATAGCGGACA
Consensus       -AA-ATG AT- --CAC-GC -C A-----A--- CAT-AAGGT- AT-GC-GA-A

                401                                                       450
SEQ ID NO: 83   TAGTCAT AAA CCACCGCG CC GGTGGAGAC C TTGAGTGGAA CCCCTTCGTG
SEQ ID NO: 85   TAGTCAT AAA CCACCGCG CC GGTGGAGGC C TCGAGTGGAA CCCCTTCGTG
SEQ ID NO: 75   TAGTCAT CAA CCACCGCG CC GGCGGCGAT C TGGAGTGGAA CCCCTTCGTG
SEQ ID NO: 77   TAGTCAT CAA CCACCGCG CC GGCGGTGAC C TGGAGTGGAA CCCCTTCGTG
SEQ ID NO: 73   TAGTCAT CAA CCACCGCG CC GGCGGCGAC C TGGAGTGGAA TCCTTTCACC
SEQ ID NO: 79   TAGTAAT CAA CCACCGCG CC GGAGGAGAC C TTGAGTGGAA CCCCTTCGTC
SEQ ID NO: 81   TCGTCAT AAA CCACCGCG CA GGCGGAGAC C TCGAGTGGAA CCCGTTCGTT
     CLONE A    TAGTGAT AAA CCACCGCG CA GGTGGAGAC C TCGAATGGAA CCCATACGTC
Consensus       T-GT-AT -AA CCACCGCG C- GG-GG-G-- C T-GA-TGGAA -CC-T-C---

                451                                                       500
SEQ ID NO: 83   AACGACT ATA CCTGGACA GA CTTCTCAAA A GTCGCCTCCG GTAAATATAC
SEQ ID NO: 85   AACGACT ATA CCTGGACA GA CTTCTCAAA A GTCGCCTCCG GTAAATATAC
SEQ ID NO: 75   AACGACT ATA CCTGGACC GA CTTCTCGAA G GTCGCGTCGG GTAAATACAC
SEQ ID NO: 77   AACGACT ATA CCTGGACC GA CTTCTCAAA G GTCGCGTCGG GTAAATACAC
SEQ ID NO: 73   AACAGCT ACA CCTGGACC GA TTTCTCGAA G GTCGCGTCGG GCAAGTACAC
SEQ ID NO: 79   AATGACT ACA CCTGGACG GA CTTCTCGAA G GTCGCTTCCG GCAAGTACAC
SEQ ID NO: 81   GGGGACT ACA CCTGGACG GA CTTCTCAAA G GTGGCCTCGG GCAAATATAC
     CLONE A    GGCGACT ATA CCTGGACG GA CTTTTCTAA G GTCGCCTCCG GGAAATACAA
Consensus       -----CT A-A CCTGGAC- GA -TT-TC-AA- GT-GC-TC-G G-AA-TA-A-
```

FIGURE 14C
(cont.)

```
                501                                                   550
SEQ ID NO: 83   GGCCAAC TAC CTTGACTT CC ACCCAAACG A GCTTCACTGT TGTGATGAAG
SEQ ID NO: 85   AGCCAAC TAC CTTGACTT CC ACCCAAACG A GCTTCACTGT TGTGATGAAG
SEQ ID NO: 75   GGCCAAC TAC CTCGACTT CC ACCCGAACG A GCTCCACGCG GGCGATTCCG
SEQ ID NO: 77   GGCCAAC TAC CTCGACTT CC ACCCGAACG A GCTCCATGCG GGCGATTCCG
SEQ ID NO: 73   GGCCAAC TAC CTCGACTT CC ACCCGAACG A GCTTCACGCG GGCGATTCCG
SEQ ID NO: 79   GGCCAAC TAC CTCGACTT CC ACCCCAACG A GGTCAAGTGC TGCGACGAGG
SEQ ID NO: 81   TGCCAAC TAC CTCGACTT CC ACCCCAACG A GGTCAAGTGC TGTGACGAGG
    CLONE A     GGCCCAC TAC ATGGACTT CC ATCCAAACA A CTACAGCACC TCAGACGAGG
Consensus       -GCC-AC TAC -T-GACTT CC A-CC-AAC- A ---------- ---GA----G

                551                                                   600
SEQ ID NO: 83   GTACCTT TGG AGGATACC CT GATATATGT C ACGACAAAAG CTGGGACCAG
SEQ ID NO: 85   GTACCTT TGG AGGATACC CT GATATATGT C ACGACAAAAG CTGGGACCAG
SEQ ID NO: 75   GAACATT TGG AGGCTATC CC GACATATGC C ACGACAAGAG CTGGGACCAG
SEQ ID NO: 77   GAACATT TGG AGGCTATC CC GACATATGC C ACGACAAGAG CTGGGACCAG
SEQ ID NO: 73   GAACATT TGG AGGCTATC CC GACATATGC C ACGACAAGAG CTGGGACCAG
SEQ ID NO: 79   GCACCTT TGG AGGGTTCC CG GACATAGCC C ACGAGAAGAG CTGGGACCAG
SEQ ID NO: 81   GCACATT TGG AGGCTTCC CA GACATAGCC C ACGAGAAGAG CTGGGACCAG
    CLONE A     GAACCTT CGG TGGCTTCC CA GACATTGAT C ACCTCGTGCC CTTCAACCAG
Consensus       G-AC-TT-GG -GG-T--CC- GA-AT---- C AC-------- CT---ACCAG

                601                                                   650
SEQ ID NO: 83   TACTGGC TCT GGGCGAGC AG CGAAAGCTA C GCTGCCTACC TCAGGAGCAT
SEQ ID NO: 85   TACTGGC TCT GGGCGAGC AG CGAAAGCTA C GCTGCCTACC TCAGGAGCAT
SEQ ID NO: 75   TACTGGC TCT GGGCCAGC CA GGAGAGCTA C GCGGCCTATC TCAGGAGCAT
SEQ ID NO: 77   TACTGGC TCT GGGCCAGC CA GGAGAGCTA C GCGGCATATC TCAGGAGCAT
SEQ ID NO: 73   CACTGGC TCT GGGCCAGC AA CGAAAGCTA C GCCGCCTACC TCCGGAGCAT
SEQ ID NO: 79   TACTGGC TCT GGGCGAGC AA CGAGAGCTA C GCCGCCTACC TCAGGAGCAT
SEQ ID NO: 81   CACTGGC TCT GGGCGAGC GA TGAGAGCTA C GCCGCCTACC TAAGGAGCAT
    CLONE A     TACTGGC TGT GGGCGAGC AA CGAGAGCTA C GCCGCCTACC TCAGGAGCAT
Consensus       -ACTGGC T-T GGGC-AGC-- -GA-AGCTA C GC-GC-TA-C T--GGAGCAT

                651                                                   700
SEQ ID NO: 83   AGGGGTT GAC GCCTGGCG TT TCGACTACG T CAAGGGCTAC GGAGCATGGG
SEQ ID NO: 85   AGGGGTT GAC GCCTGGTG TT TCGACTACG T CAAGGGCTAC GGAGCCTGGG
SEQ ID NO: 75   CGGCATC GAC GCCTGGCG CT TCGACTACG T CAAGGGCTAT GCTCCCTGGG
SEQ ID NO: 77   CGGCATC GAT GCCTGGCG CT TCGACTACG T CAAGGGCTAT GCTCCCTGGG
SEQ ID NO: 73   CGGCATC GAC GCCTGGCG CT TCGACTACG T CAAGGGCTAC GCTCCCTGGG
SEQ ID NO: 79   CGGCGTT GAC GCATGGCG CT TCGACTACG T CAAGGGCTAC GGAGCGTGGG
SEQ ID NO: 81   CGGCGTT GAT GCCTGGCG CT TTGACTACG T GAAGGGCTAC GGAGCGTGGG
    CLONE A     AGGGATC GAT GCGTGGCG CT TTGACTACG T TAAGGGCTAC GGCGCGTGGG
Consensus       -GG--T-GA- GC-TGG-G-T T-GACTACG T -AAGGGCTA- G---C-TGGG

                701                                                   750
SEQ ID NO: 83   TTGTTAA CGA CTGGCTCA GC TGGTGGGGA G GCTGGGCCGT TGGAGAGTAC
SEQ ID NO: 85   TTGTTAA CGA CTGGCTCA GC TGGTGGGGA G GCTGGGCCGT TGGAGAGTAC
SEQ ID NO: 75   TCGTCAG GGA CTGGCTGA AC TGGTGGGGA G GCTGGGCAGT TGGAGAGTAC
SEQ ID NO: 77   TCGTCAA GGA CTGGCTGA AC TGGTGGGGA G GCTGGGCGGT TGGAGAGTAC
SEQ ID NO: 73   TCGTTAA GAA CTGGCTGA AC CGGTGGGGC G GCTGGGCGGT TGGAGAGTAC
SEQ ID NO: 79   TCGTCAA GGA CTGGCTGG AC TGGTGGGGA G GCTGGGCCGT CGGGGAGTAC
SEQ ID NO: 81   TCGTCAA GGA CTGGCTCA AC TGGTGGGGC G GCTGGGCCGT TGGCGAGTAC
    CLONE A     TCGTCAA GGA CTGGCTGA GT CAGTGGGGC G GCTGGGCCGT CGGCGAGTAC
Consensus       T-GT-A- --A CTGGCT-- -- --GTGGGG- G GCTGGGC-GT -GG-GAGTAC
```

FIGURE 14C
(cont.)

```
                751                                                   800
SEQ ID NO: 83   TGGGACACGA ACGTTGATGC ACTCCTCAAC TGGGCATACA GCAGCGGCGC
SEQ ID NO: 85   TGGGACACTA ACGTTGATGC ACTCCTCAAC TGGGCATACA ACAGCGGCGC
SEQ ID NO: 75   TGGGACACCA ACGTCGACGC TGTTCTCAAC TGGGCATACT CGAGCGGTGC
SEQ ID NO: 77   TGGGACACCA ACGTCGACGC TGTTCTCAAC TGGGCATACT CGAGCGGTGC
SEQ ID NO: 73   TGGGACACCA ACGTCGATGC ACTCCTGAGC TGGGCCTACG ACAGCGGTGC
SEQ ID NO: 79   TGGGACACAA ACGTTGATGC ACTGCTCAAC TGGGCCTACT CGAGCGATGC
SEQ ID NO: 81   TGGGACACCA ACGTTGATGC ACTCCTCAAC TGGGCCTACT CGAGCGGCGC
     CLONE A    TGGGACACCA ACGTCGATGC GCTCCTCAAC TGGGCCTACA GCAGCGGCGC
Consensus       TGGGACAC-A ACGT-GA-GC --T-CT-A-C TGGGC-TAC- --AGCG--GC

                801                                                   850
SEQ ID NO: 83   CAAGGTCTTT GACTTCCCGC TCTACTACAA GATGGACGAA GCCTTCGACA
SEQ ID NO: 85   CAAGGTCTTT GACTTCCCGC TCTACTACAA GATGGACGAA GCCTTCGACA
SEQ ID NO: 75   CAAGGTCTTT GACTTCGCCC TCTACTACAA GATGGACGAG GCCTTCGATA
SEQ ID NO: 77   CAAGGTCTTT GACTTCGCCC TCTACTACAA GATGGACGAG GCCTTCGATA
SEQ ID NO: 73   TAAAGTCTTC GACTTCCCGC TCTACTACAA GATGGACGAG GCCTTCGATA
SEQ ID NO: 79   AAAAGTCTTC GACTTCCCGC TCTACTACAA GATGGACGCG GCCTTTGACA
SEQ ID NO: 81   CAAGGTCTTC GACTTCCCGC TCTACTACAA GATGGATGAG GCCTTTGACA
     CLONE A    CAAGGTCTTC GACTTCCCGC TCTACTACAA GATGGACGAG GCCTTTGACA
Consensus       -AA-GTCTT- GACTTC-C-C TCTACTACAA GATGGA-G-- GCCTT-GA-A

                851                                                   900
SEQ ID NO: 83   ACACCAACAT CCCGGCATTA GTGGATGCAC TCAGATACGG CCAGACAGTG
SEQ ID NO: 85   ATACCAACAT CCCCGCTTTG GTTTACGCCC TCAAGAATGG CGGGACAGTG
SEQ ID NO: 75   ACAACAACAT TCCCGCCCTG GTGGACGCCC TCAGATACGG CCAGACAGTG
SEQ ID NO: 77   ACAACAACAT TCCCGCCCTG GTGGACGCCC TCAGATACGG TCAGACAGTG
SEQ ID NO: 73   ACAACAACAT CCCCGCCCTC GTGGACGCCC TCAAGAACGG AGGCACGGTC
SEQ ID NO: 79   ACAAGAACAT TCCCGCACTC GTCGAGGCCC TCAAGAACGG GGGCACAGTC
SEQ ID NO: 81   ACAAAAACAT TCCAGCGCTC GTCTCTGCCC TTCAGAACGG CCAGACTGTT
     CLONE A    ACAAGAACAT TCCCGCCCTC GTTTACGCCA TCCAGAACGG TGAAACCGTC
Consensus       A-A--AACAT -CC-GC--T- GT----GC-- T-----A-GG ----AC-GT-

                901                                                   950
SEQ ID NO: 83   GTCAGCCGCG ATCCCTTCAA GGCGGTAACT TTCGTTGCCA ACCACGATAC
SEQ ID NO: 85   GTCAGCCGCG ACCCATTCAA GGCGGTAACT TTCGTTGCCA ACCACGATAC
SEQ ID NO: 75   GTCAGCCGCG ACCCGTTCAA GGCTGTGACG TTTGTAGCCA ACCACGATAC
SEQ ID NO: 77   GTCAGCCGCG ACCCGTTCAA GGCTGTGACG TTTGTAGCCA ACCACGATAC
SEQ ID NO: 73   GTCAGCCGCG ACCCGTTCAA AGCCGTGACC TTCGTTGCCA ACCACGATAC
SEQ ID NO: 79   GTCAGCCGCG ACCCGTTTAA GGCCGTAACC TTCGTTGCAA ACCACGACAC
SEQ ID NO: 81   GTCTCCCGCG ACCCGTTCAA GGCCGTAACC TTTGTAGCAA ACCACGACAC
     CLONE A    GTCAGCAGGG ATCCCTTCAA GGCCGTTACC TTCGTGGCTA ACCACGATAC
Consensus       GTC--C-G-G A-CC-TT-AA -GC-GT-AC- TT-GT-GC-A ACCACGA-AC

                951                                                  1000
SEQ ID NO: 83   AGATATAATC TGGAACAAGT ATCCGGCTTA TGCATTCATC CTTACCTATG
SEQ ID NO: 85   AGATATAATC TGGAACAAGT ATCCGGCTTA TGCATTCATC CTTACCTATG
SEQ ID NO: 75   CGACATAATC TGGAACAAGT ATCCAGCCTA CGCGTTCATC CTCACCTACG
SEQ ID NO: 77   CGACATAATC TGGAACAAGT ATCCAGCCTA CGCGTTCATC CTCACCTACG
SEQ ID NO: 73   CAACATAATC TGGAACAAGT ATCCGGCCTA CGCCTTCATC CTCACCTATG
SEQ ID NO: 79   GGACATAATT TGGAACAAGT ACCCGGCCTA CGCCTTCATC CTCACCTACG
SEQ ID NO: 81   CGATATAATC TGGAACAAGT ACCTTGCTTA TGCTTTCATC CTCACCTACG
     CLONE A    GAACATAATC TGGAACAAGT ACCCTGCCTA TGCCTTCATC CTGACCTACG
Consensus       --A-ATAAT- TGGAACAAGT A-C--GC-TA -GC-TTCATC CT-ACCTA-G
```

FIGURE 14C
(cont.)

```
                1001                                                1050
SEQ ID NO: 83   AGGGACAGCC TGTTATATTC TACCGCGACT ACGAGGAGTG GCTCAACAAG
SEQ ID NO: 85   AGGGACAGCC TGTTATATTC TACCGCGACT ACGAGGAGTG GCTCAACAAG
SEQ ID NO: 75   AGGGCCAGCC GACAATATTC TACCGCGACT ACGAGGAGTG GCTCAACAAG
SEQ ID NO: 77   AGGGCCAGCC GACAATATTC TACCGCGACT ACGAGGAGTG GCTCAACAAG
SEQ ID NO: 73   AGGGACAGCC GGCAATATTC TACCGCGACT ACGAGGAGTG GCTCAACAAG
SEQ ID NO: 79   AGGGCCAGCC GACGATATTC TACCGCGACT ACGAGGAGTG GCTCAACAAG
SEQ ID NO: 81   AAGGCCAGCC CGTCATATTT TACCGCGACT ACGAGGAGTG GCTCAACAAG
CLONE A         AAGGTCAGCC CGTCATCTTC TACCGCGACT ACGAGGAGTG GCTCAACAAG
Consensus       A-GG-CAGCC ----AT-TT- TACCGCGACT ACGAGGAGTG GCTCAACAAG

                1051                                                1100
SEQ ID NO: 83   GATAAGCTTA ACAACCTCAT CTGGATACAC GATCACCTTG CTGGAGGGAG
SEQ ID NO: 85   GATAAGCTTA ACAACCTCAT CTGGATACAC GATCACCTTG CTGGAGGGAG
SEQ ID NO: 75   GACAAGCTCA AGAACCTCAT CTGGATACAT GACAACCTCG CCGGAGGGAG
SEQ ID NO: 77   GATAAGCTCA AGAACCTCAT CTGGATACAT GACAACCTCG CCGGAGGGAG
SEQ ID NO: 73   GACAGGCTCA GGAACCTCAT CTGGATACAC GACCACCTCG CGGGAGGAAG
SEQ ID NO: 79   GACAGGCTCA AGAACCTCAT CTGGATACAC GACCACCTCG CCGGTGGAAG
SEQ ID NO: 81   GACAGGTTGA ACAACCTCAT ATGGATACAC GACCACCTCG CAGGTGGAAG
CLONE A         GACAAACTCA ACAACCTCAT ATGGATTCAC GAGCACCTGG CAGGGGGAAG
Consensus       GA-A---T-A --AACCTCAT -TGGAT-CA- GA--ACCT-G C-GG-GG-AG

                1101                                                1150
SEQ ID NO: 83   TACTGACATT GTTTACTACG ACAGCGACGA GCTTATCTTT GTGAGAAACG
SEQ ID NO: 85   TACTGACATT GTTTACTACG ACAGCGACGA GCTTATCTTT GTGAGAAACG
SEQ ID NO: 75   CACTGACATC GTTTACTACG ACAACGACGA GCTGATATTC GTGAGAAACG
SEQ ID NO: 77   CACTGACATC GTTTACTACG ACAACGACGA GCTGATATTC GTGAGAAACG
SEQ ID NO: 73   CACAGACATC ATCTACTACG ACAGCGACGA GCTTATCTTC GTGAGAAACG
SEQ ID NO: 79   CACCGACATA GTCTACTACG ATAACGATGA ACTCATCTTC GTCAGGAACG
SEQ ID NO: 81   CACGAGCATA GTTTACTACG ACAGCGACGA GATGATTTTC GTGAGGAACG
CLONE A         CACCAAGATC CTCTACTACG ACGACGATGA GCTCATCTTC ATGAGGGAAG
Consensus       -AC----AT- -T-TACTACG A---CGA-GA --T-AT-TT- -T-AG--A-G

                1151                                                1200
SEQ ID NO: 83   GCTATGGCAC CAAACCAGGA CTGATAACCT ATATCAACCT CGGCTCAAGC
SEQ ID NO: 85   GCTATGGCAC CAAACCAGGA CTGATAACCT ATATCAACCT CGGCTCAAGC
SEQ ID NO: 75   GCTACGGAAG CAAGCCGGGA CTGATAACAT ACATCAACCT CGGCTCAAGC
SEQ ID NO: 77   GCTACGGAAG CAAGCCGGGA CTGATAACAT ACATCAACCT CGCCTCAAGC
SEQ ID NO: 73   GCTACGGGGA CAAGCCGGGA CTGATAACCT ACATCAACCT CGGCTCAAGC
SEQ ID NO: 79   GCTACGGGGA CAAGCCGGGG CTTATAACCT ACATCAACCT AGGCTCGAGC
SEQ ID NO: 81   GCTATGGAAG CAAGCCTGGC CTTATAACTT ACATCAACCT CGGCTCGAGC
CLONE A         GCTACGGCGA CAGGCCCGGG CTTATAACCT ACATCAACCT CGGTAGCGAC
Consensus       GCTA-GG--- CA--CC-GG- CT-ATAAC-T A-ATCAACCT -G-------C

                1201                                                1250
SEQ ID NO: 83   AAAGTTGGAA GGTGGGTCTA CGTT...CCA AAGTTCGCCG GTTCATGCAT
SEQ ID NO: 85   AAAGCTGGAA GGTGGGTCTA CGTT...CCA AAGTTCGCCG GTTCATGCAT
SEQ ID NO: 75   AAAGCCGGAA GGTGGGTTTA CGTT...CCG AAGTTCGCAG GCTCGTGCAT
SEQ ID NO: 77   AAAGCCGGAA GGTGGGTTTA CGTT...CCG AAGTTCGCAG GCTCGTGCAT
SEQ ID NO: 73   AAGGCCGGAA GGTGGGTCTA CGTT...CCG AAGTTCGCAG GCTCGTGCAT
SEQ ID NO: 79   AAGGCCGGGA GGTGGGTCTA CGTT...CCG AAGTTCGCGG GAGCGTGCAT
SEQ ID NO: 81   AAGGTTGGAA GGTGGGTTTA TGTG...CCG AAGTTCGCGG GCGCGTGCAT
CLONE A         TGGGCGGAGA GATGGGTGAA CGTTGGCTCA AAGTTCGCGG GCTATACAAT
Consensus       ---G--G--A G-TGGGT--A -GT-----C- AAGTTCGC-G G-------AT
```

FIGURE 14C
(cont.)

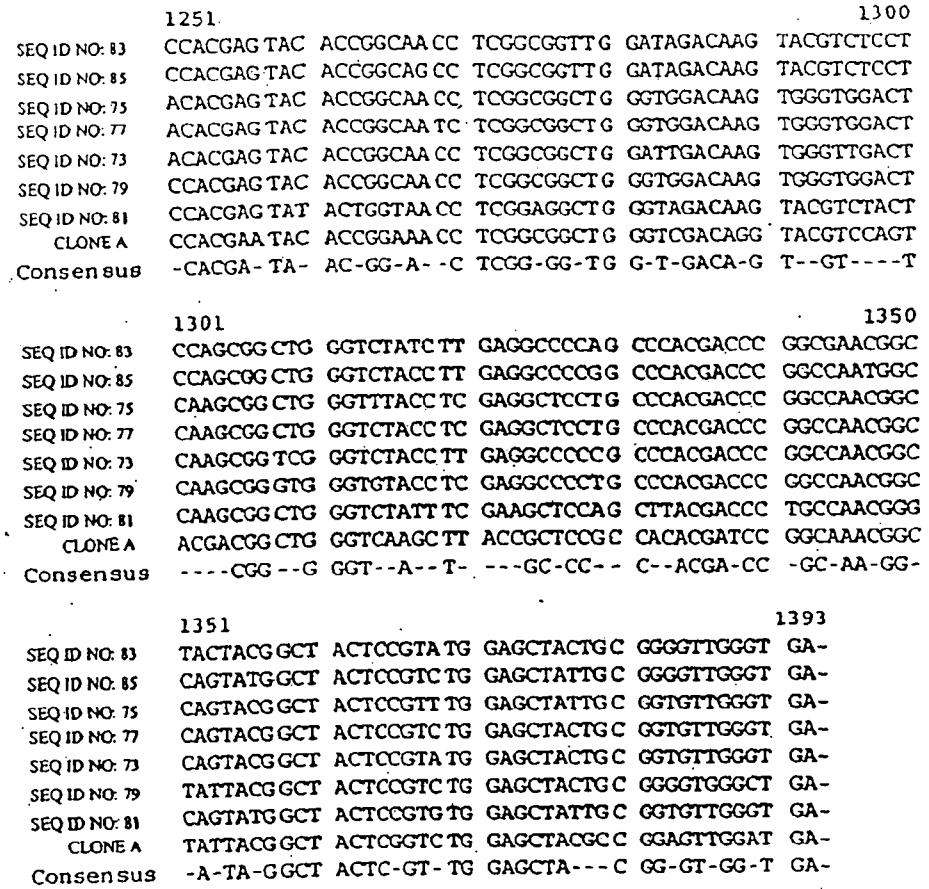
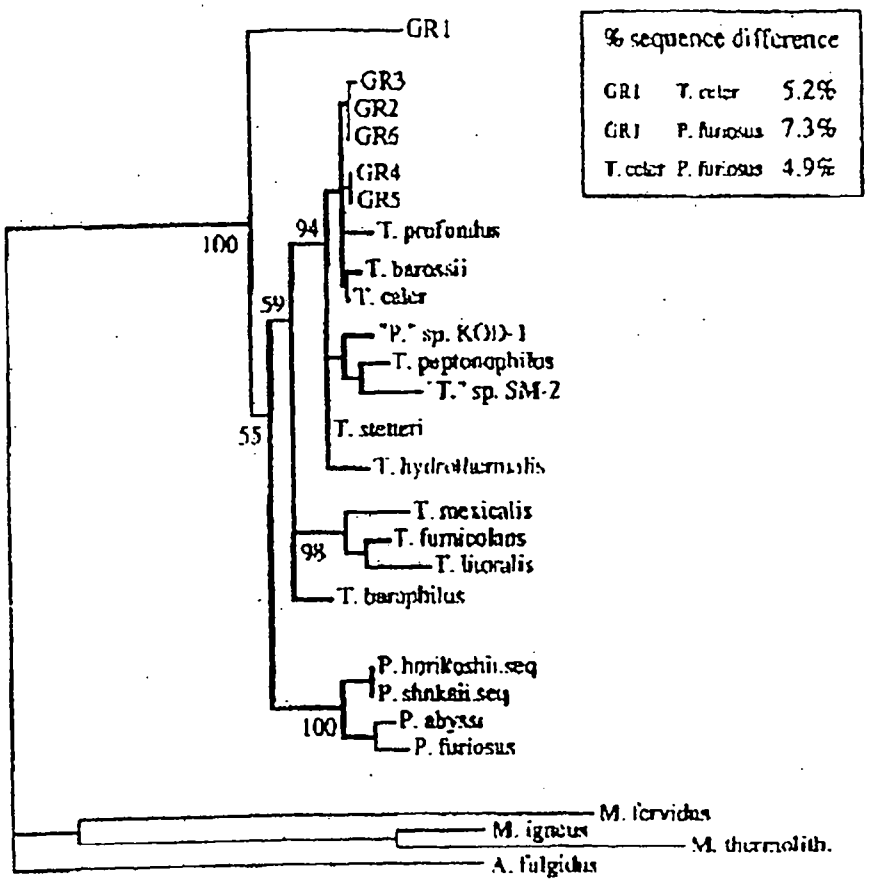


FIGURE 14C
(cont.)

Neighbor-joining tree for Thermococcales



bootstrap values for 100 replicates

Summit & Baross, Deep-Sea Research Pt. II, in press

FIGURE 15

FIGURE 16A

SEQ ID NO.: 1

atggcaaagtattccgagctcgaagagggcgggctcataatgcaggccttctactgggacgtccccatgggaggaatctggtgggacacgat
agcccagaagatacccgactgggcaagcgccgggatttcggcgatatggattccccggcgagcaagggcatgggcggcgcctattcgatg
ggctacgacccctacgacttctttgacctcggtgagtacgaccagaagggaacggtagagacgcgctttggctccaagcaggagctcgtgaa
catgataaacaccgcccacgcctatggcatgaaggtaatagccgatatagtcatcaaccaccgcgccggcggtgacctggagtggaacccctt
cgtgaacgactatacctggaccgacttctcaaaggtcgcgtcgggtaaatacacggccaactacctcgacttccacccgaacgagctccatgc
gggcgattccggaacatttggaggctatcccgacatatgccacgacaagagctgggaccagtactggctctgggccagccaggagagctac
gcggcatatctcaggagcatcggcatcgatgcctggcgcttcgactacgtcaagggctacggagcgtgggtcgtcaaggactggctggactg
gtggggaggctgggccgtcggggagtactgggacacaaacgttgatgcactgctcaactgggcctactcgagcgatgcaaaagtcttcgactt
cccgctctactacaagatggacgcggcctttgacaacaagaacattcccgcactcgtcgaggccctcaagaacgggggcacagtcgtcagcc
gcgacccgtttaaggccgtaaccttcgttgcaaaccacgacaccgatataatctggaacaagtatccagcctacgcgttcatcctcacctacgag
ggccagccgacaatattctaccgcgactacgaggagtggctcaacaaggataagctcaagaacctcatctggatacatgacaacctcgccgg
aggaagcactgacatcgtttactacgacaacgacgagctgatattcgtgagaaacggctacggaagcaagccgggactgataacatacatcaa
cctcgcctcaagcaaagccggaaggtgggtttacgttccgaagttcgcaggctcgtgcatacacgagtacaccggcaatctcggcggctgggt
ggacaagtgggtggactcaagcggctgggtctacctcgaggctcctgcccacgacccggccaacggccagtacggctactccgtctggagc
tactgcggtgttgggtga

SEQ ID NO.: 2

Met Ala Lys Tyr Ser Glu Leu Glu Glu Gly Gly Leu Ile Met Gln Ala Phe Tyr Trp Asp Val Pro Met
Gly Gly Ile Trp Trp Asp Thr Ile Ala Gln Lys Ile Pro Asp Trp Ala Ser Ala Gly Ile Ser Ala Ile Trp
Ile Pro Pro Ala Ser Lys Gly Met Gly Gly Ala Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Phe Phe Asp
Leu Gly Glu Tyr Asp Gln Lys Gly Thr Val Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Val Asn
Met Ile Asn Thr Ala His Ala Tyr Gly Met Lys Val Ile Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly
Asp Leu Glu Trp Asn Pro Phe Val Asn Asp Tyr Thr Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys
Tyr Thr Ala Asn Tyr Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe Gly Gly
Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp Leu Trp Ala Ser Gln Glu Ser Tyr Ala
Ala Tyr Leu Arg Ser Ile Gly Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp Val
Val Lys Asp Trp Leu Asp Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn Val Asp Ala
Leu Leu Asn Trp Ala Tyr Ser Ser Asp Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Ala
Ala Phe Asp Asn Lys Asn Ile Pro Ala Leu Val Glu Ala Leu Lys Asn Gly Gly Thr Val Val Ser Arg
Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala
Tyr Ala Phe Ile Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu Trp Leu Asn
Lys Asp Lys Leu Lys Asn Leu Ile Trp Ile His Asp Asn Leu Ala Gly Gly Ser Thr Asp Ile Val Tyr
Tyr Asp Asn Asp Glu Leu Ile Phe Val Arg Asn Gly Tyr Gly Ser Lys Pro Gly Leu Ile Thr Tyr Ile
Asn Leu Ala Ser Ser Lys Ala Gly Arg Trp Val Tyr Val Pro Lys Phe Ala Gly Ser Cys Ile His Glu
Tyr Thr Gly Asn Leu Gly Gly Trp Val Asp Lys Trp Val Asp Ser Ser Gly Trp Val Tyr Leu Glu Ala
Pro Ala His Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr Cys Gly Val Gly

SEQ ID NO.: 3

atggccaagtacctggagctcgaagagggcgggctcataatgcaggccttctactgggacgtccccatgggaggaatctggtgggacacgat
agcccagaagatacccgactgggcaagcgccgggatttcggcgatatggattccccggcgagcaagggcatgggcggcgcctattcgatg
ggctacgacccctacgacttctttgacctcggtgagtacgaccagaagggaacggtagagacgcgctttggctccaagcaggagctcgtgaa
catgataaacaccgcccacgcctacggcatcaaggtcatcgcagacatagtaatcaaccaccgcgccggaggagaccttgagtggaacccct
tcgtcaatgactacacctggacggacttctcgaaggtcgcttccggcaagtacacggccaattacctcgacttccacccgaacgagctccatgc
gggcgattccggaacatttggaggctatcccgacatatgccacgacaagagctgggaccagtactggctctgggccagccaggagagctac
gcggcatatctcaggagcatcggcatcgatgcctggcgcttcgactacgtcaagggctatgctccctgggtcgtcaaggactggctgaactggt
ggggaggctgggcggttggagagtactgggacaccaacgtcgacgctgttctcaactgggcatactcgagcggtgccaaggtctttgacttcg
cccctctactacaagatggatgaggcctttgacaacaaaaacattccagcgctcgtctctgcccttcagaacggccagactgttgtctcccgcgac
ccgttcaaggccgtaaccttigtagcaaaccacgacaccgatataatctggaacaagtatccagcctacgcgttcatcctcacctacgagggcc

FIGURE 16B

agccgacaatattctaccgcgactacgaggagtggctcaacaaggataagctcaagaacctcatctggatacatgacaacctcgccggagga
agcactgacatcgtttactacgacaacgacgagctgatattcgtgagaaacggctacggaagcaagccgggactgataacatacatcaacctc
gcctcaagcgaagccggaaggtgggtctacgttccgaagttcgcgggagcgtgcatccacgagtacaccggcaacctcggcggctgggtgg
acaagtgggtggactcaagcgggtgggtgtacctcgaggcccctgcccacgacccggccaacggctattacggctactccgtctggagctatt
gcggtgttgggtga

SEQ ID NO.: 4
Met Ala Lys Tyr Leu Glu Leu Glu Glu Gly Gly Leu Ile Met Gln Ala Phe Tyr Trp Asp Val Pro Met
Gly Gly Ile Trp Trp Asp Thr Ile Ala Gln Lys Ile Pro Asp Trp Ala Ser Ala Gly Ile Ser Ala Ile Trp
Ile Pro Pro Ala Ser Lys Gly Met Gly Gly Ala Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Phe Phe Asp
Leu Gly Glu Tyr Asp Gln Lys Gly Thr Val Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Val Asn
Met Ile Asn Thr Ala His Ala Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly
Asp Leu Glu Trp Asn Pro Phe Val Asn Asp Tyr Thr Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys
Tyr Thr Ala Asn Tyr Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe Gly Gly
Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp Leu Trp Ala Ser Gln Glu Ser Tyr Ala
Ala Tyr Leu Arg Ser Ile Gly Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Ala Pro Trp Val
Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn Val Asp Ala
Val Leu Asn Trp Ala Tyr Ser Ser Gly Ala Lys Val Phe Asp Phe Ala Leu Tyr Tyr Lys Met Asp Glu
Ala Phe Asp Asn Lys Asn Ile Pro Ala Leu Val Ser Ala Leu Gln Asn Gly Gln Thr Val Val Ser Arg
Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala
Tyr Ala Phe Ile Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu Trp Leu Asn
Lys Asp Lys Leu Lys Asn Leu Ile Trp Ile His Asp Asn Leu Ala Gly Gly Ser Thr Asp Ile Val Tyr
Tyr Asp Asn Asp Glu Leu Ile Phe Val Arg Asn Gly Tyr Gly Ser Lys Pro Gly Leu Ile Thr Tyr Ile
Asn Leu Ala Ser Ser Glu Ala Gly Arg Trp Val Tyr Val Pro Lys Phe Ala Gly Ala Cys Ile His Glu
Tyr Thr Gly Asn Leu Gly Gly Trp Val Asp Lys Trp Val Asp Ser Ser Gly Trp Val Tyr Leu Glu Ala
Pro Ala His Asp Pro Ala Asn Gly Tyr Tyr Gly Tyr Ser Val Trp Ser Tyr Cys Gly Val Gly

SEQ ID NO.: 5
atggccaagtactccgagctggaagagggcggcgttataatgcaggccttctactgggacgtcccaggtggaggaatctggtgggacaccat
caggagcaagataccggagtggtacgaggcgggaatatccgccatttggattcccccggcaagcaagggcatgggcggcgcctattcgatg
ggctacgacccctacgacttctttgacctcggtgagtacgaccagaagggaacggtagagacgcgctttggctccaagcaggagctcgtgaa
catgataaacaccgcccacgcctatggcatgaaggtaatagccgatatagtcataaccaccgcgccggcggtgacctggagtggaaccccctt
cgtgaacgactatacctggaccgacttctcaaaggtcgcgtcgggtaaatacacggccaactacctcgacttccacccgaacgagctccatgc
gggcgattccggaacatttggaggctatcccgacatatgccacgacaagagctgggaccagtactggctctgggccagccaggagagctac
gcggcatatctcaggagcatcggcatcgatgcctggcgcttcgactacgtcaagggctatgctccctgggtcgtcaaggactggctgaactggt
ggggaggctgggcggttggagagtactgggacaccaacgtcgacgctgttctcaactgggcatactcgagcggtgccaaggtctttgacttcg
ccctctactacaagatggatgaggccttgacaacaaaaacattccagcgctcgtctctgccctcagaacggccagactgttgtctccgcggac
ccgttcaaggccgtaaccttgtagcaaaccacgacaccgatataatctggaacaagtaccttgcttatgctttcatcctcacctacgaaggccag
cccgtcatattctaccgcgaccacgaggagtggctcaacaaggacaggttgaacaacctcatatggatacacgaccacctcgcaggtggaag
caccgacatagtctactacgataacgatgaactcatcttcgtcaggaacggctacggggacaagccggggcttataacctacatcaacctaggc
tcgagcaaggccggaaggtgggtttatgtgccgaagttcgcgggcgcgtgcatccacgagtatactggtaacctcggaggctgggtagacaa
gtacgtctactcaagcggctgggtctatctcgaagctccagcttacgaccctgccaacgggcagtatggctactccgtgtggagctactgcggg
gtgggctga

SEQ ID NO.: 6
Met Ala Lys Tyr Ser Glu Leu Glu Glu Gly Gly Val Ile Met Gln Ala Phe Tyr Trp Asp Val Pro Gly
Gly Gly Ile Trp Trp Asp Thr Ile Arg Ser Lys Ile Pro Glu Trp Tyr Glu Ala Gly Ile Ser Ala Ile Trp
Ile Pro Pro Ala Ser Lys Gly Met Gly Gly Ala Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Phe Phe Asp
Leu Gly Glu Tyr Asp Gln Lys Gly Thr Val Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Val Asn
Met Ile Asn Thr Ala His Ala Tyr Gly Met Lys Val Ile Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly

FIGURE 16C

Asp Leu Glu Trp Asn Pro Phe Val Asn Asp Tyr Thr Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe Gly Gly Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp Leu Trp Ala Ser Gln Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Ala Pro Trp Val Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn Val Asp Ala Val Leu Asn Trp Ala Tyr Ser Ser Gly Ala Lys Val Phe Asp Phe Ala Leu Tyr Tyr Lys Met Asp Glu Ala Phe Asp Asn Lys Asn Ile Pro Ala Leu Val Ser Ala Leu Gln Asn Gly Gln Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Leu Ala Tyr Ala Phe Ile Leu Thr Tyr Glu Gly Gln Pro Val Ile Phe Tyr Arg Asp His Glu Glu Trp Leu Asn Lys Asp Arg Leu Asn Asn Leu Ile Trp Ile His Asp His Leu Ala Gly Gly Ser Thr Asp Ile Val Tyr Tyr Asp Asn Asp Glu Leu Ile Phe Val Arg Asn Gly Tyr Gly Asp Lys Pro Gly Leu Ile Thr Tyr Ile Asn Leu Gly Ser Ser Lys Ala Gly Arg Trp Val Tyr Val Pro Lys Phe Ala Gly Ala Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp Val Asp Lys Tyr Val Tyr Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro Ala Tyr Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr Cys Gly Val Gly

SEQ ID NO.: 9

atggccaagtactccgagctggaagagggcgggctcataatgcaggccttctactgggacgtccccatgggaggaatctggtgggacacgat
agcccagaagataccgactgggcaagcgccgggatttcggcgatatggattcccccggcgagcaagggcatgggcggcgcctattcgatg
ggctacgacccctacgacttctttgacctcggtgagtacgaccagaagggaacggtagagacgcgctttggctccaagcaggagctcgtgaa
catgataaacacggcccatgcctacggcataaaggtcatagcggacatcgtcataaaccaccgcgcaggcggagacctcgagtggaacccg
ttcgttggggactacacctggacggacttctcaaaggtggcctcgggcaaatatactgccaactacctcgacttccacccgaacgagctccatg
cgggcgattccggaacatttggaggctatcccgacatatgccacgacaagagctgggaccagtactggctctgggccagccaggagagctac
gcggcatatctcaggagcatcggcatcgatgcctggcgcttcgactacgtcaagggctatgctccctgggtcgtcaaggactggctgaactggt
ggggaggctgggcggttggagagtactgggacaccaacgtcgacgctgttctcaactgggcatactcgagcggtgccaaggtctttgacttcg
ccctctactacaagatggacgaggccttcgataacaacaacattcccgccctggtggacgccctcagatacggtcagacagtggtcagccgcg
acccgttcaaggctgtgacgtttgtagccaaccacgataccgatataatctggaacaagtatccagcctacgcgttcatcctcacctacgagggc
cagccgacaatattctaccgcgactacgaggagtggctcaacaaggataagctcaagaacctcatctggatacatgacaacctcgccggagg
aagcactgacatcgtttactacgacaacgacgagctgatattcgcgagaaacggctacggaagcaagccgggactgataacatacatcaacct
cgcctcaagcaaagccggaaggtgggtttacgttccgaagttcgcaggctcgtgcatacacgagtacaccggcaatctcggcggctgggtgg
acaagtgggtggactcaagcggctgggtctacctcgaggctcctgcccacgacccggccaacggccagtacggctactccgtctggagctac
tgcggtgttgggtga

SEQ ID NO.: 10

Met Ala Lys Tyr Ser Glu Leu Glu Glu Gly Gly Leu Ile Met Gln Ala Phe Tyr Trp Asp Val Pro Met Gly Gly Ile Trp Trp Asp Thr Ile Ala Gln Lys Ile Pro Asp Trp Ala Ser Ala Gly Ile Ser Ala Ile Trp Ile Pro Pro Ala Ser Lys Gly Met Gly Gly Ala Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Phe Phe Asp Leu Gly Glu Tyr Asp Gln Lys Gly Thr Val Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Val Asn Met Ile Asn Thr Ala His Ala Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Gly Asp Tyr Thr Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe Gly Gly Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp Leu Trp Ala Ser Gln Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Ala Pro Trp Val Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn Val Asp Ala Val Leu Asn Trp Ala Tyr Ser Ser Gly Ala Lys Val Phe Asp Phe Ala Leu Tyr Tyr Lys Met Asp Glu Ala Phe Asp Asn Asn Asn Ile Pro Ala Leu Val Asp Ala Leu Arg Tyr Gly Gln Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu Trp Leu Asn Lys Asp Lys Leu Lys Asn Leu Ile Trp Ile His Asp Asn Leu Ala Gly Gly Ser Thr Asp Ile Val Tyr Tyr Asp Asn Asp Glu Leu Ile Phe Ala Arg Asn Gly Tyr Gly Ser Lys Pro Gly Leu Ile Thr Tyr Ile Asn Leu Ala Ser Ser Lys Ala Gly Arg Trp Val Tyr Val Pro Lys Phe Ala Gly Ser Cys Ile His Glu

FIGURE 16D

Tyr Thr Gly Asn Leu Gly Gly Trp Val Asp Lys Trp Val Asp Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro Ala His Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr Cys Gly Val Gly

SEQ ID NO.: 11

atggccaagtacctggagctcgaggagggcgggctcataatgcaggccttctactgggacgtccccatgggaggaatctggtgggacacgat
agcccagaagataccgactgggcaagcgccgggatttcggcgatatggattcccccggcgagcaagggcatgggcggcgcctattcgatg
ggctacgacccctacgacttctttgacctcggtgagtacgaccagaagggaacggtagagacgcgctttggctccaagcaggagctcgtgaa
catgataaacaccgcccacgcctatggcatgaaggtaatagccgatatagtcatcaaccaccgcgccggcggtgacctggagtggaaccccctt
cgtgaacgactatacctggaccgacttctcaaaggtcgcgtcgggtaaatacacggccaactacctcgacttccacccgaacgagctccatgc
gggcgattccggaacatttggaggctatcccgacatatgccacgacaagagctgggaccagtactggctctgggccagccaggagagctac
gcggcatatctcaggagcatcggcatcgatgcctggcgcttcgactacgtcaagggctatgctccctgggtcgtcaaggactggctgaactggt
ggggaggctgggcggttggagagtactgggacaccaacgtcgacgctgttctcaactgggcatactcgagcggtgccaaggtctttgacttcg
ccctctactacaagatggacgaggccttcgataacaacaacattcccgccctggtggacgccctcagatacggtcagacagtggtcagccgcg
acccgttcaaggctgtgacgtttgtagccaaccacgataccgatataatctggaacaagtatccagcctacgcgttcatcctcacctacgagggc
cagccgacaatattctaccgcgactacgaggagtggctcaacaaggatacgctcaagaacctcatctggatacatgacaacctcgccggagg
aagcacgagcatagtttactacgacagcgacgagatgatcttcgtgaggaacggctatggaagcaagcctggccttataacttacatcaacctc
ggctcgagcaaggttggaaggtgggtctacgttccgaagttcgcgggagcgtgcatccacgagtacaccggcaacctcggcggctgggtgg
acaagtgggtggactcaagcgggtgggtgtacctcgaggcccctgcccacgacccggccaacggctattacggctactccgtctggagctac
tgcggtgttggctga

SEQ ID NO.: 12

Met Ala Lys Tyr Leu Glu Leu Glu Glu Gly Gly Leu Ile Met Gln Ala Phe Tyr Trp Asp Val Pro Met Gly Gly Ile Trp Trp Asp Thr Ile Ala Gln Lys Ile Pro Asp Trp Ala Ser Ala Gly Ile Ser Ala Ile Trp Ile Pro Pro Ala Ser Lys Gly Met Gly Gly Ala Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Phe Phe Asp Leu Gly Glu Tyr Asp Gln Lys Gly Thr Val Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Val Asn Met Ile Asn Thr Ala His Ala Tyr Gly Met Lys Val Ile Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Asn Asp Tyr Thr Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe Gly Gly Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp Leu Trp Ala Ser Gln Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Ala Pro Trp Val Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn Val Asp Ala Val Leu Asn Trp Ala Tyr Ser Ser Gly Ala Lys Val Phe Asp Phe Ala Leu Tyr Tyr Lys Met Asp Glu Ala Phe Asp Asn Asn Asn Ile Pro Ala Leu Val Asp Ala Leu Arg Tyr Gly Gln Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu Trp Leu Asn Lys Asp Thr Leu Lys Asn Leu Ile Trp Ile His Asp Asn Leu Ala Gly Gly Ser Thr Ser Ile Val Tyr Tyr Asp Ser Asp Glu Met Ile Phe Val Arg Asn Gly Tyr Gly Ser Lys Pro Gly Leu Ile Thr Tyr Ile Asn Leu Gly Ser Ser Lys Val Gly Arg Trp Val Tyr Val Pro Lys Phe Ala Gly Ala Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp Val Asp Lys Trp Val Asp Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro Ala His Asp Pro Ala Asn Gly Tyr Tyr Gly Tyr Ser Val Trp Ser Tyr Cys Gly Val Gly

SEQ ID NO.: 13

atggccaagtacctggagctcgaagagggcggggtcataatgcaggcgttctactgggacgtgccttcaggaggaatatggtgggacacaat
acggcagaagataccggagtggtacgatgccggaatctccgcaatatggattccccggcgagcaagggcatgggcggcgcctattcgatg
ggctacgacccctacgacttctttgacctcggtgagtatgaccagaagggaacggtagagacgcgctttggctccaagcaggagctcgtgaac
atgataaacacggcacatgcctacggcataaaggtcatagcggacatcgtcataaaccaccgcgcaggcggagacctcgagtggaacccgtt
cgttggggactacacctggacggacttctcaaaggtggcctcgggcaaatatactgccaactacctcgacttccaccccaacgaggtcaagtg
ctgtgacgagggcacatttggaggcttcccagacatagcccacgagaagagctgggaccagcactggctctgggcgagcgatgagagctac
gccgcctacctaaggagcatcggcgttgatgcctggcgcttcgactacgtcaagggctacggagcgtgggtcgtcaaggactggctggactg
gtggggaggctgggccgtcggggagtactgggacacaaacgttgatgcactgctcaacctgggcctactcgagcgatgcacaagtcttcgactt

FIGURE 16E

cccgctctactacaagatggatgaggcctttgacaacaaaaacattccagcgctcgtctctgcccttcagaacggccagactgttgtctcccgcg
acccgttcaaggccgtaaccttttgtagcaaaccacgacaccgatataatctggaacaagtatccagcctacgcgttcatcctcacctacgaggg
ccagccgacaatattctaccgcgactacgaggagtggctcaacaaggataagctcaagaacctcatctggatacatgacaacctcgccggag
gaagcactgacatagtctactacgataacgatgaactcatcttcgtcaggaacggctacggggacaagccggggcttataacctacatcaacct
aggctcgagcaaggccggaaggtgggtttatgtgccgaagttcgcgggcgcgtgcatccacgagtatactggtaacctcggaggctgggtag
acaagtacgtctactcaagcggctgggtctatctcgaagctccagcttacgaccctgccaacgggcagtatggctactccgtgtggagctactg
cggtgttggctga

SEQ ID NO.: 14
Met Ala Lys Tyr Leu Glu Leu Glu Glu Gly Gly Val Ile Met Gln Ala Phe Tyr Trp Asp Val Pro Ser
Gly Gly Ile Trp Trp Asp Thr Ile Arg Gln Lys Ile Pro Glu Trp Tyr Asp Ala Gly Ile Ser Ala Ile Trp
Ile Pro Pro Ala Ser Lys Gly Met Gly Gly Ala Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Phe Phe Asp
Leu Gly Glu Tyr Asp Gln Lys Gly Thr Val Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Val Asn
Met Ile Asn Thr Ala His Ala Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly
Asp Leu Glu Trp Asn Pro Phe Val Gly Asp Tyr Thr Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys
Tyr Thr Ala Asn Tyr Leu Asp Phe His Pro Asn Glu Val Lys Cys Cys Asp Glu Gly Thr Phe Gly
Gly Phe Pro Asp Ile Ala His Glu Lys Ser Trp Asp Gln His Trp Leu Trp Ala Ser Asp Glu Ser Tyr
Ala Ala Tyr Leu Arg Ser Ile Gly Val Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp
Val Val Lys Asp Trp Leu Asp Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn Val Asp
Ala Leu Leu Asn Trp Ala Tyr Ser Ser Asp Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp
Glu Ala Phe Asp Asn Lys Asn Ile Pro Ala Leu Val Ser Ala Leu Gln Asn Gly Gln Thr Val Val Ser
Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro
Ala Tyr Ala Phe Ile Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu Trp Leu
Asn Lys Asp Lys Leu Lys Asn Leu Ile Trp Ile His Asp Asn Leu Ala Gly Gly Ser Thr Asp Ile Val
Tyr Tyr Asp Asn Asp Glu Leu Ile Phe Val Arg Asn Gly Tyr Gly Asp Lys Pro Gly Leu Ile Thr Tyr
Ile Asn Leu Gly Ser Ser Lys Ala Gly Arg Trp Val Tyr Val Pro Lys Phe Ala Gly Ala Cys Ile His
Glu Tyr Thr Gly Asn Leu Gly Gly Trp Val Asp Lys Tyr Val Tyr Ser Ser Gly Trp Val Tyr Leu Glu
Ala Pro Ala Tyr Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr Cys Gly Val Gly

SEQ ID NO.: 15
atggccaagtactccgagctggaagagggcgggctcataatgcaggccttctactgggacgtccccatgggaggaatctggtgggacacgat
agcccagaagataccgactgggcaagcgccgggatttcggcgatatggattcccccggcgagcaagggcatgggcggcgcctattcgatg
ggctacgacccctacgacttctttgacctcggtgagtacgaccagaagggaacggtagagacgcgctttggctccaagcaggagctcgtgaa
catgataaacacggcccatgcctacggcataaaggtcatagcggacatcgtcataaaccaccgcgcaggcggagacctcgagtggaacccg
ttcgttggggactacacctggacggacttctcaaaggtggcctcgggcaaatatactgccaactacctcgacttccacccgaacgagctccatg
cgggcgattccggaacatttggaggctatccgacatatgccacgacaagagctgggaccagtactggctctgggccagccaggagagctac
gcgggcatatctcaggagcatcggcctcgatgcctggcgcttttgactacgtcaagggctacggagcgtggtggtcgtcaagtactggctggactg
gtgggggaggctgggccgtcgggagtactgggacacaaacgttgatgcactgctcaactgggcctactcgagcgatgcaaaagtcttcgactt
cccgctctactacaagatggatgaggcctttgacaacaaaaacattccagcgctcgtctctgcccttcagaacggccagactgttgtctcccgcg
acccgttcaaggccgtaacctttgtagcaaaccacgacaccgatataatttggaacaagtacccggcctacgccttcatcctcacctacgaggg
ccagccgacgatattctaccgcgactacgaggagtggctcaacaaggacaggctcaagaacctcatctggatacacgaccaccttgccggtg
gaagcactgacatcgtttactacgacaacgacgagctgatattcgtgagaaacggctacggaagcaagccgggactgataacatacatcaacc
tcgcctcaagcaaagccggaaggtgggtttatgtgccgaagttcgcgggcgcgtgcatccacgagtatactggtaacctcggaggctgggtag
acaagtacgtctactcaagcggctgggtctatctcgaagctccagcttacgaccctgccaacgggcagtatggctactccgtgtggagctattgc
ggtgttgggtga

SEQ ID NO.: 16
Met Ala Lys Tyr Ser Glu Leu Glu Glu Gly Gly Leu Ile Met Gln Ala Phe Tyr Trp Asp Val Pro Met
Gly Gly Ile Trp Trp Asp Thr Ile Ala Gln Lys Ile Pro Asp Trp Ala Ser Ala Gly Ile Ser Ala Ile Trp
Ile Pro Pro Ala Ser Lys Gly Met Gly Gly Ala Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Phe Phe Asp

FIGURE 16F

Leu Gly Glu Tyr Asp Gln Lys Gly Thr Val Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Val Asn
Met Ile Asn Thr Ala His Ala Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly
Asp Leu Glu Trp Asn Pro Phe Val Gly Asp Tyr Thr Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys
Tyr Thr Ala Asn Tyr Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe Gly Gly
Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp Leu Trp Ala Ser Gln Glu Ser Tyr Ala
Ala Tyr Leu Arg Ser Ile Gly Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp Val
Val Lys Asp Trp Leu Asp Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn Val Asp Ala
Leu Leu Asn Trp Ala Tyr Ser Ser Asp Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Glu
Ala Phe Asp Asn Lys Asn Ile Pro Ala Leu Val Ser Ala Leu Gln Asn Gly Gln Thr Val Val Ser Arg
Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala
Tyr Ala Phe Ile Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu Trp Leu Asn
Lys Asp Arg Leu Lys Asn Leu Ile Trp Ile His Asp His Leu Ala Gly Gly Ser Thr Asp Ile Val Tyr
Tyr Asp Asn Asp Glu Leu Ile Phe Val Arg Asn Gly Tyr Gly Ser Lys Pro Gly Leu Ile Thr Tyr Ile
Asn Leu Ala Ser Ser Lys Ala Gly Arg Trp Val Tyr Val Pro Lys Phe Ala Gly Ala Cys Ile His Glu
Tyr Thr Gly Asn Leu Gly Gly Trp Val Asp Lys Tyr Val Tyr Ser Ser Gly Trp Val Tyr Leu Glu Ala
Pro Ala Tyr Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr Cys Gly Val Gly

SEQ ID NO.: 17

atggccaagtactccgagctggaaggggcgggctcataatgcaggccttctactgggacgtccccatgggaggaatctggtgggacacgat
agcccagaagatacccgactgggcaagcgccgggatttcggcgatatggattccccggcgagcaagggcatgggcggcgcctattcgatg
ggctacgaccctacgacttctttgacctcggtgagtacgaccaggagggaacggtagagacgcgctttggctccaagcaggagctcgtgaa
catgataaacacggcccatgcctacggcataaaggtcatagcggacatcgtcataaccaccgcgcaggcggagacctcgagtggaacccg
ttcgttggggactacacctggacggacttctcaaaggtggcctcgggcaaatatactgccaactacctcgacttccaccccaacgaggtcaagt
gctgtgacgagggcacatttggaggcttcccagacatagcccacgagaagagctgggaccagcactggctctgggcgagcgatgagagcta
cgccgcctacctaaggagcatcggcgttgatgcctggcgcttcgactacgtcaagggctacggagcgtgggtcgtcaaggactggctggact
ggtggggaggctgggccgtcggggagtactgggacacaaacgttgatgcactgctcaactgggcctactcgagcgatgcaaaagtcttcgac
ttcccgctctactacaagatggacgcggccttgacaacaagaacattcccgcactcgtcgaggccctcaagaacgggggcacagtcgtcagc
cgcgacccgtttaaggccgtaaccttcgttgcaaaccacgacaccgatataatctggaacaagtatccagcctacgcgttcatcctcacctacga
gggccagccgacaatattctaccgcgactacgaggagtggctcaacaaggataagctcaagaacctcatctggatacatgacaacctcgccg
gaggaagcacgagcatagtttactacgacagcgacgagatgatcttcgtgaggaacggctatggaagcaagcctggccttataacttacatcaa
cctcggctcgagcaaggttggaaggtgggtttacgttccgaagttcgcaggctcgtgcatacacgagtacaccggcaatctcggcggctgggt
ggacaagtgggtggactcaagcggctgggtctacctcgaggctcctgcccacgacccggccaacggccagtacggctactccgtctggagc
tactgcggtgttgggtga

SEQ ID NO.: 18

Met Ala Lys Tyr Ser Glu Leu Glu Gly Gly Gly Leu Ile Met Gln Ala Phe Tyr Trp Asp Val Pro Met
Gly Gly Ile Trp Trp Asp Thr Ile Ala Gln Lys Ile Pro Asp Trp Ala Ser Ala Gly Ile Ser Ala Ile Trp
Ile Pro Pro Ala Ser Lys Gly Met Gly Gly Ala Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Phe Phe Asp
Leu Gly Glu Tyr Asp Gln Glu Gly Thr Val Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Val Asn
Met Ile Asn Thr Ala His Ala Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly
Asp Leu Glu Trp Asn Pro Phe Val Gly Asp Tyr Thr Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys
Tyr Thr Ala Asn Tyr Leu Asp Phe His Pro Asn Glu Val Lys Cys Cys Asp Glu Gly Thr Phe Gly
Gly Phe Pro Asp Ile Ala His Glu Lys Ser Trp Asp Gln His Trp Leu Trp Ala Ser Asp Glu Ser Tyr
Ala Ala Tyr Leu Arg Ser Ile Gly Val Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp
Val Val Lys Asp Trp Leu Asp Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn Val Asp
Ala Leu Leu Asn Trp Ala Tyr Ser Ser Asp Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp
Ala Ala Phe Asp Asn Lys Asn Ile Pro Ala Leu Val Glu Ala Leu Lys Asn Gly Gly Thr Val Val Ser
Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro
Ala Tyr Ala Phe Ile Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu Trp Leu
Asn Lys Asp Lys Leu Lys Asn Leu Ile Trp Ile His Asp Asn Leu Ala Gly Gly Ser Thr Ser Ile Val

FIGURE 16G

Tyr Tyr Asp Ser Asp Glu Met Ile Phe Val Arg Asn Gly Tyr Gly Ser Lys Pro Gly Leu Ile Thr Tyr Ile Asn Leu Gly Ser Ser Lys Val Gly Arg Trp Val Tyr Val Pro Lys Phe Ala Gly Ser Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp Val Asp Lys Trp Val Asp Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro Ala His Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr Cys Gly Val Gly

SEQ ID NO.: 19

atggccaagtacctggagctcgaagagggcgggctcataatgcaggccttctactgggacgtccccatgggaggaatctggtgggacacgat
agcccagaagatacccgactgggcaagcgccgggatttcggcgatatggattcctcccgcgagcaagggtatgagcggcggctattcgatgg
gctacgacccctacgattattttgaccttggtgagtactaccagaagggaacggtggaaacgaggttcggctcaaagcaggagctcataaacat
gataaacacggcccatgcctacggcataaaggtcatagcggacatcgtcataaaccacccgcgcaggcggagacctcgagtggaacccgttc
gttggggactacacctggacggacttctcaaaggtggcctcgggcaaatatactgccaactacctcgacttccacccgaacgagctccatgcg
ggcgattccggaacatttggaggctatcccgacatatgccacgacaagagctgggaccagtactggctctgggccagccaggagagctacgc
ggcatatctcaggagcatcggcatcgatgcctggcgcttcgactacgtcaagggctatgctccctgggtcgtcaaggactggctgaactggtgg
gggggctgggcggttggagagtactgggacaccaacgtcgacgctgttctcaactgggcatactcgagcggtgccaaggtctttgacttcgcc
ctctactacaagatggatgaggcctttgacaacaaaaacattccagcgctcgtctctgcccttcagaacggccagactgttgtctcccgcgaccc
gttcaaggccgtaacctttgtagcaaaccacgacaccgatataatttggaacaagtacccggcctacgccttcatcctcacctacgagggccag
ccgacgatattctaccgcgactacgaggagtggctcaacaaggacaggctcaagaacctcatctggatacacgaccacctcgccggtggaag
cactgacatcgtttactacgacaacgacgagctgatattcgtgagaaacggctacggaagcaagccgggactgataacatacatcaacctcgc
ctcaagcaaagccggaaggtgggtttatgtgccgaagttcgcgggcgcgtgcatccacgagcatactggtaacctcggaggctgggtagaca
agtacgtctactcaagcggctgggtctatctcgaagctccagcttacgaccctgccaacgggcagtatggctactccgtgtggagctactgcgg
tgttggctga

SEQ ID NO.: 20

Met Ala Lys Tyr Leu Glu Leu Glu Glu Gly Gly Leu Ile Met Gln Ala Phe Tyr Trp Asp Val Pro Met Gly Gly Ile Trp Trp Asp Thr Ile Ala Gln Lys Ile Pro Asp Trp Ala Ser Ala Gly Ile Ser Ala Ile Trp Ile Pro Pro Ala Ser Lys Gly Met Ser Gly Gly Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Tyr Phe Asp Leu Gly Glu Tyr Tyr Gln Lys Gly Thr Val Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Ile Asn Met Ile Asn Thr Ala His Ala Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Gly Asp Tyr Thr Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe Gly Gly Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp Leu Trp Ala Ser Gln Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Ala Pro Trp Val Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn Val Asp Ala Val Leu Asn Trp Ala Tyr Ser Ser Gly Ala Lys Val Phe Asp Phe Ala Leu Tyr Tyr Lys Met Asp Glu Ala Phe Asp Asn Lys Asn Ile Pro Ala Leu Val Ser Ala Leu Gln Asn Gly Gln Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu Trp Leu Asn Lys Asp Arg Leu Lys Asn Leu Ile Trp Ile His Asp His Leu Ala Gly Gly Ser Thr Asp Ile Val Tyr Tyr Asp Asn Asp Glu Leu Ile Phe Val Arg Asn Gly Tyr Gly Ser Lys Pro Gly Leu Ile Thr Tyr Ile Asn Leu Ala Ser Ser Lys Ala Gly Arg Trp Val Tyr Val Pro Lys Phe Ala Gly Ala Cys Ile His Glu His Thr Gly Asn Leu Gly Gly Trp Val Asp Lys Tyr Val Tyr Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro Ala Tyr Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr Cys Gly Val Gly

SEQ ID NO.: 21

atggccaagtactccgagctggaagagggcggcgttataatgcaggccttctactgggacgtcccaggtggaggaatctggtgggacaccat
caggagcaagataccggagtggtacgaggcgggaatatccgccatttggattcctcccgggagcaagggtatgagcggcggctattcgatgg
gctacgacccctacgatgatttggacctgggtgagtactaccagaagggaacggtggaaacgaggttcggctcaaagcaggagctcataaac
atgataaacacggcccatgcctacggcataaaggtcatagcggacatcgtcataaaccacccgcgcaggcggagacctcgagtggaacccgtt
cgttggggactacacctggacggacttctcaaaggtggcctcgggcaaatatactgccaactacctcgacttccacccgaacgagctccatgc
gggcgattccggaacatttggaggctatcccgacatatgccacgacaagagctgggaccagtactggctctgggccagccaggagagctac

FIGURE 16H

gcggtatatctcaggagcatcggcatcgatgcctggcgcttcgactacgtcaagggctacggagcgtgggtcgtcaaggactggctggactg
gtggggaggctgggccgtcggggagtactgggacacaaacgttgatgcactgctcaactgggcctactcgagcgatgcaaaagtcttcgactt
cccgctctactacaagatggatgaggcctttgacaacaaaaacattccagcgctcgtctctgcccttcagaacggccagactgttgtctcccgcg
acccgttcaaggccgtaacctttgtagcaaaccacgacaccgatataatttggaacaagtacccggcctacgccttcatcctcacctacgaggg
ccagccgacgatattctaccgcgactacgaggagtggctcaacaaggacaggctcaagaacctcatctggatacacgactacctcgccggtg
gaagcactgacatcgtttactacgacaacgacgagctgatattcgtgagaaacggctacggaagcaagccgggactgataacatacatcaacc
tcgcctcaagcaaagccggaaggtgggtttatgtgccgaagttcgcgggcgcgtgcatccacgagtatactggtaacctcggaggctgggtag
acaagtacgtctactcaagcggctgggtctatctcgaagctccagcttacgaccctgccaacgggcagtatggctactccgtgtggagctattgc
ggtgttggctga

SEQ ID NO.: 22
Met Ala Lys Tyr Ser Glu Leu Glu Glu Gly Gly Val Ile Met Gln Ala Phe Tyr Trp Asp Val Pro Gly
Gly Gly Ile Trp Trp Asp Thr Ile Arg Ser Lys Ile Pro Glu Trp Tyr Glu Ala Gly Ile Ser Ala Ile Trp
Ile Pro Pro Gly Ser Lys Gly Met Ser Gly Gly Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Asp Leu Asp
Leu Gly Glu Tyr Tyr Gln Lys Gly Thr Val Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Ile Asn Met
Ile Asn Thr Ala His Ala Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly Asp
Leu Glu Trp Asn Pro Phe Val Gly Asp Tyr Thr Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr
Thr Ala Asn Tyr Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe Gly Gly Tyr
Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp Leu Trp Ala Ser Gln Glu Ser Tyr Ala Val
Tyr Leu Arg Ser Ile Gly Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp Val Val
Lys Asp Trp Leu Asp Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn Val Asp Ala
Leu Leu Asn Trp Ala Tyr Ser Ser Asp Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Glu
Ala Phe Asp Asn Lys Asn Ile Pro Ala Leu Val Ser Ala Leu Gln Asn Gly Gln Thr Val Val Ser Arg
Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala
Tyr Ala Phe Ile Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu Trp Leu Asn
Lys Asp Arg Leu Lys Asn Leu Ile Trp Ile His Asp Tyr Leu Ala Gly Gly Ser Thr Asp Ile Val Tyr
Tyr Asp Asn Asp Glu Leu Ile Phe Val Arg Asn Gly Tyr Gly Ser Lys Pro Gly Leu Ile Thr Tyr Ile
Asn Leu Ala Ser Ser Lys Ala Gly Arg Trp Val Tyr Val Pro Lys Phe Ala Gly Ala Cys Ile His Glu
Tyr Thr Gly Asn Leu Gly Gly Trp Val Asp Lys Tyr Val Tyr Ser Ser Gly Trp Val Tyr Leu Glu Ala
Pro Ala Tyr Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr Cys Gly Val Gly

SEQ ID NO.: 23
atggccaagtactccgagctggaagagggcggcgttatagtgcaggccttctactgggacgtcccaggtggaggaatctggtgggacaccat
caggagcaagataccggagtggtacgaggcgggaatatccgccatttggattcccccggcgagcaagggcatgggcggcgcctattcgatg
ggctacgacccctacgacttctttgacctcggtgagtacgaccagaagggaacggtagagacgcgctttggctccaagcaggagctcgtgaa
catgataaacacggccatgcctacggcataaaggtcatagcggacatcgtcataaaccacggcgcaggcggagacctcgagtggaaccg
ttcgttggggactacaactggacggacttctcaaaggtggcctcgggcaaatatactgccaactacctcgacttccacccgaacgagctccatg
cgggcgattccggaacatttggaggctatcccgacatatgccacgacaagagctgggaccagtactggctctgggccagccaggagagctac
gcggcatatctcaggagcatcggcatcgatgcctggcgcttcgactacgtcaagggctacggagcgtgggtcgtcaaggactggctggactg
gtggggaggctgggccgtcggggagtactgggacacaaacgttgatgcactgctcaactgggcctactcgagcgatgcaaaagtcttcgactt
cccgctctactacaagatggatgaggcctttgacaacaaaaacattccagcgctcgtctctgcccttcagaacggccagactgttgtctcccgcg
acccgttcaaggccgtaacctttgtagcaaaccacgacaccgatataatctggaacaagtatccagcctacgcgttcatcctcacctacgaggg
ccagccgacaatattctaccgcgactacgaggagtggctcaacaaggataagctcaagaacctcatctggatacatgacaacctcgccggag
gaagcatgagcatagtttactacgacagcgacgagatgatcttcgtgaggaacggctatggaagcaagcctggccttataacttacatcaacctc
ggctcgagcaaggttggaaggtgggtctacgttccgaagttcgcgggagcgtgcatccacgagtacaccggcaacctcggcggctgggtgg
acaagtgggtggactcaagcgggtgggtgtacctcgaggcccctgcccacgacccggccaacggctattacggctactccgtctggagctatt
gcggtgttggctga

SEQ ID NO.: 24

FIGURE 16I

Met Ala Lys Tyr Ser Glu Leu Glu Glu Gly Gly Val Ile Val Gln Ala Phe Tyr Trp Asp Val Pro Gly Gly Gly Ile Trp Trp Asp Thr Ile Arg Ser Lys Ile Pro Glu Trp Tyr Glu Ala Gly Ile Ser Ala Ile Trp Ile Pro Pro Ala Ser Lys Gly Met Gly Gly Ala Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Phe Phe Asp Leu Gly Glu Tyr Asp Gln Lys Gly Thr Val Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Val Asn Met Ile Asn Thr Ala His Ala Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Gly Asp Tyr Thr Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe Gly Gly Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp Leu Trp Ala Ser Gln Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp Val Val Lys Asp Trp Leu Asp Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn Val Asp Ala Leu Leu Asn Trp Ala Tyr Ser Ser Asp Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Glu Ala Phe Asp Asn Lys Asn Ile Pro Ala Leu Val Ser Ala Leu Gln Asn Gly Gln Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu Trp Leu Asn Lys Asp Lys Leu Lys Asn Leu Ile Trp Ile His Asp Asn Leu Ala Gly Gly Ser Met Ser Ile Val Tyr Tyr Asp Ser Asp Glu Met Ile Phe Val Arg Asn Gly Tyr Gly Ser Lys Pro Gly Leu Ile Thr Tyr Ile Asn Leu Gly Ser Ser Lys Val Gly Arg Trp Val Tyr Val Pro Lys Phe Ala Gly Ala Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp Val Asp Lys Trp Val Asp Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro Ala His Asp Pro Ala Asn Gly Tyr Tyr Gly Tyr Ser Val Trp Ser Tyr Cys Gly Val Gly

SEQ ID NO.: 25

atggccaagtacctggagctcgaagagggcgggctcataatgcaggccttctactgggacgtccccatgggaggaatctggtgggacacgat
agcccagaagataccegactgggcaagcgccgggatttcggcgatatggattcctcccgcgagcaagggtatgagcggcggctattcgatgg
gctacgaccctacgattattttgacctcggtgagtactaccagaagggaacggtggaaacgaggttcggctcaaagcaggagctcataaacat
gataaacaccgcccacgcctatggcatgaaggtaatagccgatatagtcatcaaccaccgcgccggcggtgacctggagtggaaccccttcgt
gaacgactatacctggaccgacttctcaaaggtcgcgtcgggtaaatacacggccaactacctcgacttccacccgaacgagctccatgcggg
cgattccggaacatttggaggctatcccgacatatgccacgacaagagctgggaccagtactggctctgggccagccaggagagctacgcgg
catatctcaggagcatcggcatcgatgcctggcgcttcgactacgtcaagggctatgctccctgggtcgtcaaggactggctgaactggtgggg
aggctgggcggttggagagtactgggacaccaacgtcgacgctgttctcaactgggcatactcgagcggtgccaaggtctttgacttcgccctc
tactacaagatggacgaggccttcgataacaacaacattcccgcccctggtgggcgccctcagatacggtcagacagtggtcagccgcgaccc
gttcaaggctgtgacgtttgtagccaaccacgatacccgatataatctggaacaagtatccagcctacgcgttcatcctcacctacgagggccagc
cgacaatattctaccgcgactacgaggagtggctcaacaaggataagctcaagaacctcatctggatacatgacaacctcgccggaggaagc
accgacatagtctactacgataacgatgaactcatcttcgtcaggcacggctacggggacaagccggggcttataacctacatcaacctaggct
cgagcaaggccggaaggtgggtttacgttccgaagttcgcaggctcgtgcatacacgagtacaccggcaatctcggcggctgggtggacaa
gtgggtggactcaagcggctgggtctacctcgaggctcctgcccacgacccggccaacggccagtacggctactccgtctggagctattgcg
gtgttgggtga

SEQ ID NO.: 26

Met Ala Lys Tyr Leu Glu Leu Glu Glu Gly Gly Leu Ile Met Gln Ala Phe Tyr Trp Asp Val Pro Met Gly Gly Ile Trp Trp Asp Thr Ile Ala Gln Lys Ile Pro Asp Trp Ala Ser Ala Gly Ile Ser Ala Ile Trp Ile Pro Pro Ala Ser Lys Gly Met Ser Gly Gly Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Tyr Phe Asp Leu Gly Glu Tyr Tyr Gln Lys Gly Thr Val Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Ile Asn Met Ile Asn Thr Ala His Ala Tyr Gly Met Lys Val Ile Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Asn Asp Tyr Thr Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe Gly Gly Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp Leu Trp Ala Ser Gln Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Ala Pro Trp Val Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn Val Asp Ala Val Leu Asn Trp Ala Tyr Ser Ser Gly Ala Lys Val Phe Asp Phe Ala Leu Tyr Tyr Lys Met Asp Glu Ala Phe Asp Asn Asn Asn Ile Pro Ala Leu Val Gly Ala Leu Arg Tyr Gly Gln Thr Val Val Ser Arg Asp

FIGURE 16J

Pro Phe Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu Trp Leu Asn Lys Asp Lys Leu Lys Asn Leu Ile Trp Ile His Asp Asn Leu Ala Gly Gly Ser Thr Asp Ile Val Tyr Tyr Asp Asn Asp Glu Leu Ile Phe Val Arg His Gly Tyr Gly Asp Lys Pro Gly Leu Ile Thr Tyr Ile Asn Leu Gly Ser Ser Lys Ala Gly Arg Trp Val Tyr Val Pro Lys Phe Ala Gly Ser Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp Val Asp Lys Trp Val Asp Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro Ala His Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr Cys Gly Val Gly

SEQ ID NO.: 27

atggcaaagtattccgagctcgaagagggcggcgttataatgcaggccttctactgggacgtcccaggtggaggaatctggtgggacaccatc
aggagcaagataccggagtggtacgaggcgggaatatccgccatttggattcctcccgcgagcaagggtatgagcggcggctattcgatggg
ctacgacccctacgattattttgacctcggtgagtactaccagaagggaacggtggaaacgaggttcggctcaaagcaggagctcataaacatg
ataaacacggcccatgcctacggcataaaggtcatagcggacatcgtcataaaccaccgcgcaggcggagacctcgagtggaacccgttcgt
tggggactacacctggacggacttctcaaaggtggcctcgggcaaatatactgccaactacctcgacttccacccgaacgagctccatgcggg
cgattccggaacatttggaggctatcccgacatatgccacgacaagagctgggaccagtactggctctgggccagccaggagagctacgcgg
catatctcaggagcatcggcatcgatgcctggcgcttcgactacgtcaagggctatgctccctgggtcgtcaaggactggctgaactggtgggg
aggctgggcggttggagagtactgggacaccaacgtcgacgctgttctcaactgggcatactcgagcggtgccaaggtctttgacttcgccctc
tactacaagatggacgcggcctttgacaacaagaacattcccgcactcgtcgaggccctcaagaacgggggcacagtcgtcagccgcgacc
cgtttaaggccgtaaccttcgttgcaaaccacgacaccgatataatctggaacaagtatccagcctacgcgttcatcctcacctacgagggccag
ccgacaatattctaccgcgactacgaggagtggctcaacaaggataagctcaagaacctcatctggatacatgacaacctcgccggaggaag
cactgacatcgtttactacgacaacgacgagctgatattcgtgagaaacggctacggaagcaagccgggactgataacatacatcaacctcgc
gtcaagcaaagccggaaggtgggtttacgttccgaagttcgcaggctcgtgcatacacgagtacaccggcaatctcggcggctgggtggaca
agtgggtggactcaagcggctgggtctacctcgaggctcctgcccacgacccggccaacggccagtacggctactccgtctggagctactgc
ggtgttgggtga

SEQ ID NO.: 28

Met Ala Lys Tyr Ser Glu Leu Glu Glu Gly Gly Val Ile Met Gln Ala Phe Tyr Trp Asp Val Pro Gly Gly Gly Ile Trp Trp Asp Thr Ile Arg Ser Lys Ile Pro Glu Trp Tyr Glu Ala Gly Ile Ser Ala Ile Trp Ile Pro Pro Ala Ser Lys Gly Met Ser Gly Gly Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Tyr Phe Asp Leu Gly Glu Tyr Tyr Gln Lys Gly Thr Val Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Ile Asn Met Ile Asn Thr Ala His Ala Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Gly Asp Tyr Thr Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe Gly Gly Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp Leu Trp Ala Ser Gln Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Ala Pro Trp Val Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn Val Asp Ala Val Leu Asn Trp Ala Tyr Ser Ser Gly Ala Lys Val Phe Asp Phe Ala Leu Tyr Tyr Lys Met Asp Ala Ala Phe Asp Asn Lys Asn Ile Pro Ala Leu Val Glu Ala Leu Lys Asn Gly Gly Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu Trp Leu Asn Lys Asp Lys Leu Lys Asn Leu Ile Trp Ile His Asp Asn Leu Ala Gly Gly Ser Thr Asp Ile Val Tyr Tyr Asp Asn Asp Glu Leu Ile Phe Val Arg Asn Gly Tyr Gly Ser Lys Pro Gly Leu Ile Thr Tyr Ile Asn Leu Ala Ser Ser Lys Ala Gly Arg Trp Val Tyr Val Pro Lys Phe Ala Gly Ser Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp Val Asp Lys Trp Val Asp Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro Ala His Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr Cys Gly Val Gly

SEQ ID NO.: 29

atggccaagtacctggagctcgaagagggcgggctcataatgcaggccttctactgggacgtccccatgggaggaatctggtgggacacggt
agcccagaagataccccgactgggcaagcgccgggatttcggcgatatggattcccccggcgagcaagggcatgggcggcgcctattcgatg
ggctacgacccctacgacttctttgacctcggtgagtacgaccagaagggaacggtagagacgcgcttggctccaagcaggagctcgtgaa

FIGURE 16K

catgataaacacggcccatgcctacggcataaaggtcatagcggacatcgtcataaaccaccgcgcaggcggagacctcgagtggaacccg
ttcgttggggactacacctggacggacttctcaaaggtggtctcgggcaaatatactgccaactacctcgacttccacccgaacgagctccatgc
gggcgattccggaacatttggaggctatcccgacatatgccacgacaagagctgggaccagtactggctctgggccagccaggagagctac
gcggcatatctcaggagcatcggcatcgatgcctggcgcttcgactacgtcaagggctatgctccctgggtcgtcaaggactggctgaactggt
ggggaggctgggcggttggagagtactgggacaccaacgtcgacgctgttctcaactgggcatactcgagcggtgccaaggtctttgacttcg
ccctctactacaagatggatgaggcctttgacaacaaaaacattccagcgctcgtctctgcccttcagaacggccagactgttgtctcccgcgac
ccgttcaaggccgtaacctttgtagcaaaccacgacaccgatataatctggaacaagtaccttgcttatgccttcatcctcacctacgaaggccag
cccgtcatattctaccgcgactacgaggagtggctcaacaaggacaggttgaacaacctcatatggatacacgaccacctcgcagggggaag
caccgacatagtctactacgataacgatgaactcatcttcgtcaggaacggctacggggacaagccggggcttataacctacatcaacctaggc
tcgagcaaggccggaaggtgggtttatgtgccgaagttcgcgggcgcgtgcatccacgagtatactggtaacctcggaggctgggtagacaa
gtacgtctactcaagcggctgggtctatctcgaagctccagcttacgaccctgccaacgggcagtatggctactccgtgtggagctactgcggt
gttgggtga

SEQ ID NO.: 30
Met Ala Lys Tyr Leu Glu Leu Glu Glu Gly Gly Leu Ile Met Gln Ala Phe Tyr Trp Asp Val Pro Met
Gly Gly Ile Trp Trp Asp Thr Val Ala Gln Lys Ile Pro Asp Trp Ala Ser Ala Gly Ile Ser Ala Ile Trp
Ile Pro Pro Ala Ser Lys Gly Met Gly Gly Ala Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Phe Phe Asp
Leu Gly Glu Tyr Asp Gln Lys Gly Thr Val Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Val Asn
Met Ile Asn Thr Ala His Ala Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly
Asp Leu Glu Trp Asn Pro Phe Val Gly Asp Tyr Thr Trp Thr Asp Phe Ser Lys Val Val Ser Gly Lys
Tyr Thr Ala Asn Tyr Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe Gly Gly
Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp Leu Trp Ala Ser Gln Glu Ser Tyr Ala
Ala Tyr Leu Arg Ser Ile Gly Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Ala Pro Trp Val
Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn Val Asp Ala
Val Leu Asn Trp Ala Tyr Ser Ser Gly Ala Lys Val Phe Asp Phe Ala Leu Tyr Tyr Lys Met Asp Glu
Ala Phe Asp Asn Lys Asn Ile Pro Ala Leu Val Ser Ala Leu Gln Asn Gly Gln Thr Val Val Ser Arg
Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Leu Ala
Tyr Ala Phe Ile Leu Thr Tyr Glu Gly Gln Pro Val Ile Phe Tyr Arg Asp Tyr Glu Glu Trp Leu Asn
Lys Asp Arg Leu Asn Asn Leu Ile Trp Ile His Asp His Leu Ala Gly Gly Ser Thr Asp Ile Val Tyr
Tyr Asp Asn Asp Glu Leu Ile Phe Val Arg Asn Gly Tyr Gly Asp Lys Pro Gly Leu Ile Thr Tyr Ile
Asn Leu Gly Ser Ser Lys Ala Gly Arg Trp Val Tyr Val Pro Lys Phe Ala Gly Ala Cys Ile His Glu
Tyr Thr Gly Asn Leu Gly Gly Trp Val Asp Lys Tyr Val Tyr Ser Ser Gly Trp Val Tyr Leu Glu Ala
Pro Ala Tyr Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr Cys Gly Val Gly

SEQ ID NO.: 31
atggcaaagtactccgagctggaagagggcggcgttataatgcaggccttctactgggacgtcccatggtggaggaatctggtgggacaccat
caggagcaggataccggagtggtacgaggcgggaatatccgccatttggattccctgggcgagcaagggcatgggcgggcgcctatttgatg
ggctacgaccctacgacttcttgacctcggtgagtacgaccagaagggaacggtagagacgcgctttggctccaagcaggagctcgtgaa
catgataaacacggcccatgcctacggcataaaggtcatagcggacatcgtcataaaccaccgcgcaggcggagacctcgagtggaacccg
ttcgttggggactacacctggacggacttctcaaaggtggcctcgggcaaatatactgccaactacctcgacttccacccgaacgagctccatg
cgggcgattccggaacatttggaggctatcccgacatatgccacgacaagagctgggaccagtactggctctgggccagccaggagagctac
gcggcatatctcaggagcatcggcatcgatgcctggcgctttgactacgtgaagggctacggagcgtgggtcgtcaaggactggctcaactgg
tggggcggctgggccgttggcgagtactgggacaccaacgttgatgcactcctcaactgggcctactcgagcggcgccaaggtcttcgacttc
ccgctctactacaagatggacgaggccttcgataacaacaacattcccgccctggtggacgccctcagatacggtcagacagtggtcagccgc
gacccgttcaaggctgtgacgtttgtagccaaccacgatacccgatataatctggaacaagtatccagcctacgcgttcatcctcacctacgaggg
ccagccgacaatattctaccgcgactacgaggagtggctcaacaaggataagctcaagaacctcatctggatacatgacaacctggccggag
gaagcacgagcatagtttactacgacagcgacgagatgatcttcgtgaggaccggctatggaagcaagcctggccttataacttacatcaacct
cggctcgagcaaggttggaaggtgggtttatgtgccgaagttcgcgggcgcgtgcatccacgagtatactggtaacctcggaggctgggtaga
caagtacgtctactcaagcggctgggtctatctcgaagctccagcttacgaccctgccaacgggcagtatggctactccgtgtggagctattgcg
gtgttggctga

FIGURE 16L

SEQ ID NO.: 32
Met Ala Lys Tyr Ser Glu Leu Glu Glu Gly Gly Val Ile Met Gln Ala Phe Tyr Trp Asp Val Pro Gly Gly Gly Ile Trp Trp Asp Thr Ile Arg Ser Arg Ile Pro Glu Trp Tyr Glu Ala Gly Ile Ser Ala Ile Trp Ile Pro Pro Ala Ser Lys Gly Met Gly Gly Ala Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Phe Phe Asp Leu Gly Glu Tyr Asp Gln Lys Gly Thr Val Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Val Asn Met Ile Asn Thr Ala His Ala Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Gly Asp Tyr Thr Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe Gly Gly Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp Leu Trp Ala Ser Gln Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp Val Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn Val Asp Ala Leu Leu Asn Trp Ala Tyr Ser Ser Gly Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Glu Ala Phe Asp Asn Asn Asn Ile Pro Ala Leu Val Asp Ala Leu Arg Tyr Gly Gln Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu Trp Leu Asn Lys Asp Lys Leu Lys Asn Leu Ile Trp Ile His Asp Asn Leu Ala Gly Gly Ser Thr Ser Ile Val Tyr Tyr Asp Ser Asp Glu Met Ile Phe Val Arg Thr Gly Tyr Gly Ser Lys Pro Gly Leu Ile Thr Tyr Ile Asn Leu Gly Ser Ser Lys Val Gly Arg Trp Val Tyr Val Pro Lys Phe Ala Gly Ala Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp Val Asp Lys Tyr Val Tyr Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro Ala Tyr Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr Cys Gly Val Gly

SEQ ID NO.: 33
atggccaagtactccgagctggaagagggcggggtcataatgcaggcgttctactgggacgtgccttcaggaggaatatggtgggacacaat
acggcagaagataccggagtggtacgatgccggaatctccgcaatatggattcctcccgcgagcaagggtatgagcggcggctattcgatgg
gctacgacccctacgattattttgacctcggtgagtactaccagaagggaacggtggaaacgaggttcggctcaaagcaggagctcataaacat
gataaacacggcccatgcctacggcataaaggtcatagcggacatcgtcataaaccaccgcgcaggcggagacctcgagtggaacccgttc
gttggggactacacctggacggacttctcaaaggtggcctcgggcaaatatactgccaactacctcgacttccacccgaacgagctccatgcg
ggcgattccggaacatttggaggctatcccgacatatgccacgacaagagctgggaccagtactggctctgggccagccaggagagctacgc
ggcatatctcaggagcatcggcatcgatgcctggcgctttgactacgtgaagggctacggagcgtgggtcgtcaaggactggctcaactggtg
gggcggctgggccgttggcgagtactgggacaccaacgttgatgcactcctcaactgggcctactcgagcggcgccaaggtcttcgactttcc
gctctactacaagatggacgcggcctttgacaacaagaacattcccgcactcgtcgaggccctcaagaacgggggcacagtcgtcagccgcg
acccgtttaaggccgtaaccttcgttgcaaaccacgacaccgatataatctggaccaagtaccttgcttatgctttcatcctcacctacgaaggcca
gcccgtcatattctaccgcgactacgaggagtggctcaacaaggacaggttgaacaacctcatatggatacacgaccacctcgcaggtggaag
caccgacatagtctactacgataacgatgaactcatcttcgtcaggaacggctacggggacaagccggggcttataacctacatcaacctaggc
tcgagcaaggccggaaggtgggtttacgttccgaagttcgcaggctcgtgcatacacgagtacaccgggcaatctcggcggctggtggacaa
atggcgtcgactcaagcggctgggtctacctcgaggctcctgccacgaccggccaacggccagtacgggtactccgtctggagctactgcg
gtgttggctga

SEQ ID NO.: 34
Met Ala Lys Tyr Ser Glu Leu Glu Glu Gly Gly Val Ile Met Gln Ala Phe Tyr Trp Asp Val Pro Ser Gly Gly Ile Trp Trp Asp Thr Ile Arg Gln Lys Ile Pro Glu Trp Tyr Asp Ala Gly Ile Ser Ala Ile Trp Ile Pro Pro Ala Ser Lys Gly Met Ser Gly Gly Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Tyr Phe Asp Leu Gly Glu Tyr Tyr Gln Lys Gly Thr Val Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Ile Asn Met Ile Asn Thr Ala His Ala Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Gly Asp Tyr Thr Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe Gly Gly Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp Leu Trp Ala Ser Gln Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp Val Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn Val Asp Ala

FIGURE 16M

Leu Leu Asn Trp Ala Tyr Ser Ser Gly Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Ala
Ala Phe Asp Asn Lys Asn Ile Pro Ala Leu Val Glu Ala Leu Lys Asn Gly Gly Thr Val Val Ser Arg
Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Asp Ile Ile Trp Thr Lys Tyr Leu Ala
Tyr Ala Phe Ile Leu Thr Tyr Glu Gly Gln Pro Val Ile Phe Tyr Arg Asp Tyr Glu Glu Trp Leu Asn
Lys Asp Arg Leu Asn Asn Leu Ile Trp Ile His Asp His Leu Ala Gly Gly Ser Thr Asp Ile Val Tyr
Tyr Asp Asn Asp Glu Leu Ile Phe Val Arg Asn Gly Tyr Gly Asp Lys Pro Gly Leu Ile Thr Tyr Ile
Asn Leu Gly Ser Ser Lys Ala Gly Arg Trp Val Tyr Val Pro Lys Phe Ala Gly Ser Cys Ile His Glu
Tyr Thr Gly Asn Leu Gly Gly Trp Val Asp Lys Trp Val Asp Ser Ser Gly Trp Val Tyr Leu Glu Ala
Pro Ala His Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr Cys Gly Val Gly

SEQ ID NO.: 35

atggccaagtactccgagctggaagagggcggcgttataatgcaggccttctactgggacgtcccaggtggaggaatctggtgggacaccat
caggagcaagataccggagtggtacgaggcgggaatatccgccatttggattcccccggcgagcaagggcatgggcggcgcctattcgatg
ggctacgacccctacgacttctttgacctcggtgagtacgaccagaagggaacggtagagacgcgctttggctccaagcaggagctcgtgaa
catgataaacaccgcccacgcctacggcatcaaggtcatcgcagacatagtaatcaaccaccgcgccggaggagaccttgagtggaacccct
tcgtcaatgactacacctggacggacttctcgaaggtcgcttccggcaagtacacggccaactacctcgacttccaccccaacgaggtcaagtg
ctgtgacgagggcacatttggaggcttcccagacatagcccacgagaagagctgggaccagcactggctctgggcgagcgatgagagctac
gccgcctacctaaggagcatcggcgttgatgcctggcgcttcgactacgtcaagggctatgctccctgggtcgtcaaggactggctgaactggt
ggggaggctgggcggttggagagtactgggacaccaacgtcgacgctgttctcaactgggcatactcgagcggtgccaaggtctttgacttcg
ccctctactacaagatggacgcggcctttgacaacaagaacattcccgcactcgtcgaggccctcaagaacgggggcacagtcgtcagccgc
gacccgtttaaggccgtaaccttcgttgcaaaccacgacaccgatataatctggaacaagtatccagcctacgcgttcatcctcacctacgaggg
ccagccgacaatattctaccgcgactacgaggagtggctcaacaaggataagctcaagaacctcatctggatacatgacaacgtcgccggag
gaagcaccgacatagtctactacgataacgatgaactcatcttcgtcaggaacggctacggggacaagccggggcttataacctacatcaacct
aggctcgagcaaggccggaaggtgggtttacgttccgaagttcgcaggctcgtgcatacacgagtacaccggcaatctcggcggctgggtgg
acaagtgggtggactcaagcggctgggtctacctcgaggctcctgcccacgacccggccaacggccagtacggctactccgtctggagctac
tgcggtgttgggtga

SEQ ID NO.: 36

Met Ala Lys Tyr Ser Glu Leu Glu Glu Gly Gly Val Ile Met Gln Ala Phe Tyr Trp Asp Val Pro Gly
Gly Gly Ile Trp Trp Asp Thr Ile Arg Ser Lys Ile Pro Glu Trp Tyr Glu Ala Gly Ile Ser Ala Ile Trp
Ile Pro Pro Ala Ser Lys Gly Met Gly Gly Ala Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Phe Phe Asp
Leu Gly Glu Tyr Asp Gln Lys Gly Thr Val Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Val Asn
Met Ile Asn Thr Ala His Ala Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly
Asp Leu Glu Trp Asn Pro Phe Val Asn Asp Tyr Thr Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys
Tyr Thr Ala Asn Tyr Leu Asp Phe His Pro Asn Glu Val Lys Cys Cys Asp Glu Gly Thr Phe Gly
Gly Phe Pro Asp Ile Ala His Glu Lys Ser Trp Asp Gln His Trp Leu Trp Ala Ser Asp Glu Ser Tyr
Ala Ala Tyr Leu Arg Ser Ile Gly Val Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Ala Pro Trp
Val Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn Val Asp
Ala Val Leu Asn Trp Ala Tyr Ser Ser Gly Ala Lys Val Phe Asp Phe Ala Leu Tyr Tyr Lys Met Asp
Ala Ala Phe Asp Asn Lys Asn Ile Pro Ala Leu Val Glu Ala Leu Lys Asn Gly Gly Thr Val Val Ser
Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro
Ala Tyr Ala Phe Ile Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu Trp Leu
Asn Lys Asp Lys Leu Lys Asn Leu Ile Trp Ile His Asp Asn Val Ala Gly Gly Ser Thr Asp Ile Val
Tyr Tyr Asp Asn Asp Glu Leu Ile Phe Val Arg Asn Gly Tyr Gly Asp Lys Pro Gly Leu Ile Thr Tyr
Ile Asn Leu Gly Ser Ser Lys Ala Gly Arg Trp Val Tyr Val Pro Lys Phe Ala Gly Ser Cys Ile His
Glu Tyr Thr Gly Asn Leu Gly Gly Trp Val Asp Lys Trp Val Asp Ser Ser Gly Trp Val Tyr Leu Glu
Ala Pro Ala His Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr Cys Gly Val Gly

SEQ ID NO.: 71

atggccaagtacctggagctcgaagagggcggggtcataatgcaggcgttctactgggacgtgccttcaggaggaatatggtgggacacaat

FIGURE 16N

acggcagaagataccggagtggtacgatgccggaatctccgcaatatggattccccccggcgagcaagggcatgggcggcgcctattcgatg
ggctacgacccctacgacttctttgacctcggtgagtacgaccagaagggaacggtagagacgcgctttggctccaagcaggagctcgtgaa
catgataaacacggcccatgcctacggcataaaggtcatagcggacatcgtcataaaccaccgcgcaggcggagacctcgagtggaacccg
ttcgttggggactacacctggacggacttctcaaaggtagcctcgggcaaatatactgccaactacctcgacttccacccgaacgagctccatg
cgggcgattccggaacatttggaggctatcccgacatatgccacgacaagagctgggaccagtactggctctgggccagccaggagagctac
gcggcatatctcaggagcatcggcatcgatgcctggcgcttcgactacgtcaagggctatgctccctgggtcgtcaaggactggctgaactggt
ggggaggctgggcggttggagagtactgggacaccaacgtcgacgctgttctcaactgggcatactcgagcggtgccaaggtctttgacttcg
ccctctactacaagatggatgaggcctttgacaacaaaaacattccagcgctcgtctctgcccttcagaacggccagactgttgtctcccgcgac
ccgttcaaggccgtaacctttgtagcaaaccacgacaccgatataatctggaacaagtatccagcctacgcgttcatcctcacctacgagggcc
agccgacaatattctaccgcgactacgaggagtggctcaacaaggataagctcaagaacctcatctggatacatgacaacctcgccggagga
agcactgacatcgtttactacgacaacgacgagctgatattcgtgagaaacggctacggaagcaagccgggactgataacatacatcaacctc
gcctcaagcaaagccggaaggtgggtttatgtgccgaagttcgcgggcgcgtgcatccacgagtatactggtaacctcggaggctgggtaga
caagtacgtctactcaagcggctgggtctatctcgaagctccagcttacgaccctgccaacgggcagtatggctactccgtgtggagctactgc
ggggtgggctga

SEQ ID NO.: 72
Met Ala Lys Tyr Leu Glu Leu Glu Glu Gly Gly Val Ile Met Gln Ala Phe Tyr Trp Asp Val Pro Ser
Gly Gly Ile Trp Trp Asp Thr Ile Arg Gln Lys Ile Pro Glu Trp Tyr Asp Ala Gly Ile Ser Ala Ile Trp
Ile Pro Pro Ala Ser Lys Gly Met Gly Gly Ala Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Phe Phe Asp
Leu Gly Glu Tyr Asp Gln Lys Gly Thr Val Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Val Asn
Met Ile Asn Thr Ala His Ala Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly
Asp Leu Glu Trp Asn Pro Phe Val Gly Asp Tyr Thr Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys
Tyr Thr Ala Asn Tyr Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe Gly Gly
Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp Leu Trp Ala Ser Gln Glu Ser Tyr Ala
Ala Tyr Leu Arg Ser Ile Gly Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Ala Pro Trp Val
Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn Val Asp Ala
Val Leu Asn Trp Ala Tyr Ser Ser Gly Ala Lys Val Phe Asp Phe Ala Leu Tyr Tyr Lys Met Asp Glu
Ala Phe Asp Asn Lys Asn Ile Pro Ala Leu Val Ser Ala Leu Gln Asn Gly Gln Thr Val Val Ser Arg
Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala
Tyr Ala Phe Ile Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu Trp Leu Asn
Lys Asp Lys Leu Lys Asn Leu Ile Trp Ile His Asp Asn Leu Ala Gly Gly Ser Thr Asp Ile Val Tyr
Tyr Asp Asn Asp Glu Leu Ile Phe Val Arg Asn Gly Tyr Gly Ser Lys Pro Gly Leu Ile Thr Tyr Ile
Asn Leu Ala Ser Ser Lys Ala Gly Arg Trp Val Tyr Val Pro Lys Phe Ala Gly Ala Cys Ile His Glu
Tyr Thr Gly Asn Leu Gly Gly Trp Val Asp Lys Tyr Val Tyr Ser Ser Gly Trp Val Tyr Leu Glu Ala
Pro Ala Tyr Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr Cys Gly Val Gly

SEQ ID NO.: 49
gtggtttatgacgatgtccgctatgaccttatgccgtaggcatgggccgtgtttatcatgttcacgagctcctgcttggagccaaagcgcgtctct
accgttcccttctggtcgtactcaccgaggtcaaagaagtgtagggggtcgtagcccatcgaataggcgccgcccatgcccttgctcgccggg
ggaatccatatcgccgaaatcccggcgcttgcccagtcgggtatcttctgggctatcgtgtcccaccagattcctcccatggggacgtcccagta
gaaggcctgcattatgagcccgccctcttcgagcccggaatactttgccataagttacctcctactagtagattaaaattctgtttcctgtgtgaaatt
gtt

SEQ ID NO.: 50
Val Val Tyr Asp Asp Val Arg Tyr Asp Leu Tyr Ala Val Gly Met Gly Arg Val Tyr His Val His Glu
Leu Leu Leu Gly Ala Lys Ala Arg Leu Tyr Arg Ser Leu Leu Val Val Leu Thr Glu Val Lys Glu Val
Val Gly Val Val Ala His Arg Ile Gly Ala Ala His Ala Leu Ala Arg Arg Gly Asn Pro Tyr Arg Arg
Asn Pro Gly Ala Cys Pro Val Gly Tyr Leu Leu Gly Tyr Arg Val Pro Pro Asp Ser Ser His Gly Asp
Val Pro Val Glu Gly Leu His Tyr Glu Pro Ala Leu Phe Glu Pro Gly Ile Leu Cys His Lys Leu Pro
Pro Thr Ser Arg Leu Lys Phe Cys Phe Leu Cys Glu Ile Val

FIGURE 16O

SEQ ID NO.: 51
ATGGCCAAGTACCTGGAGCTCGAAGAGGGCGGGGTCATAATGCAGGCGTTCTACTGGG
ACGTGCCTTCAGGAGGAATATGGTGGGACACAATACGGCAGAAGATACCGGAGTGGT
ACGATGCCGGAATCTCCGCAATATGGATTCCCCCGGCGAGCAAGGGCATGGGCGGCGC
CTATTCGATGGGCTACGACCCCTACGACTTCTTTGACCTCGGTGAGTACGACCAGAAG
GGAACGGTAGAGACGCGCTTTGGCTCCAAGCAGGAGCTCGTGAACATGATAAACACC
GCCCACGCCTATGGCATGAAGGTAATAGCCGATATAGTCATCAACCACCGCGCCGGCG
GTGACCTGGAGTGGAACCCCTTCGTGAACGACTATACCTGGACCGACTTCTCAAAGGT
CGCGTCGGGTAAATACACGGCCAACTACCTCGACTTCCACCCCAACGAGGTCAAGTGC
TGTGACGAGGGCACATTTGGAGGCTTCCCAGACATAGCCCACGAGAAGAGCTGGGAC
CAGCACTGGCTCTGGGCGAGCGATGAGAGCTACGCCGCCTACCTAAGGAGCATCGGCG
TTGATGCCTGGCGCTTTGACTACGTGAAGGGCTACGGAGCGTGGGTCGTCAAGGACTG
GCTCAACTGGTGGGGCGGCTGGGCCGTTGGCGAGTACTGGGACACCAACGTTGATGCA
CTCCTCAACTGGGCCTACTCGAGCGGCGCCAAGGTCTTCGACTTCCCGCTCTACTACAA
GATGGATGAGGCCTTTGACAACAAAAACATTCCAGCGCTCGTCTCTGCCCTTCAGAAC
GGCCAGACTGTTGTCTCCCGCGACCCGTTCAAGGCCGTAACCTTTGTAGCAAACCACG
ACACCGATATAATCTGGAACAAGTATCCAGCCTACGCGTTCATCCTCACCTACGAGGG
CCAGCCGACAATATTCTACCGCGACTACGAGGAGTGGCTCAACAAGGATAAGCTCAAG
AACCTCATCTGGATACATGACAACCTCGCCGGAGGAAGCACTGACATCGTTTACTACG
ACAACGACGAGCTGATATTCGTGAGAAACGGCTACGGAAGCAAGCCGGGACTGATAA
CATACATCAACCTCGCCTCAAGCAAAGCCGGAAGGTGGGTTTACGTTCCGAAGTTCGC
AGGCTCGTGCATACACGAGTACACCGGCAATCTCGGCGGCTGGGTGGACAAGTGGGTG
GACTCAAGCGGCTGGGTCTACCTCGAGGCTCCTGCCCACGACCCGGCCAACGGCCAGT
ACGGCTACTCCGTCTGGAGCTATTGCGGTGTTGGCTGA

SEQ ID NO.: 52
MAKYLELEEGGVIMQAFYWDVPSGGIWWDTIRQKIPEWYDAGISAIWIPPASKGMGGAYS
MGYDPYDFFDLGEYDQKGTVETRFGSKQELVNMINTAHAYGMKVIADIVINHRAGGDLE
WNPFVNDYTWTDFSKVASGKYTANYLDFHPNEVKCCDEGTFGGFPDIAHEKSWDQHWL
WASDESYAAYLRSIGVDAWRFDYVKGYGAWVVKDWLNWWGGWAVGEYWDTNVDAL
LNWAYSSGAKVFDFPLYYKMDEAFDNKNIPALVSALQNGQTVVSRDPFKAVTFVANHDT
DIIWNKYPAYAFILTYEGQPTIFYRDYEEWLNKDKLKNLIWIHDNLAGGSTDIVYYDNDELI
FVRNGYGSKPGLITYINLASSKAGRWVYVPKFAGSCIHEYTGNLGGWVDKWVDSSGWVY
LEAPAHDPANGQYGYSVWSYCGVG

SEQ ID NO.: 37
atggccaagtacctggagctcgaagagggcggggtcataatgcaggcgttctactgggacgtgccttcaggaggaatatggtgggacacaat
acggcagaagataccggagtggtacgatgccggaatctccgcaatatggattcccccggcgagcaagggcatgggcggcgcctattcgatg
ggctacgacccctacgacttctttgacctcggtgagtacgaccagaagggaacggtagagacgcgctttggctccaagcaggagctcgtgaa
catgataaacaccgcccacgcctatggcatgaaggtaatagccgatatagtcatcaaccaccgcgccggcggtgacctggagtggaacccctt
cgtgaacgactataccctggaccgacttctcaaaggtcgcgtcgggtaaatacacggccaactacctcgacttccaccccgaacgagctccatgc
gggcgattccggaacatttggaggctatcccgacatatgccacgacaagagctgggaccagtactggctctgggccagcaggagagctac
gcggcatatctcaggagcatcggcatcgatgcctggcgctttgactacgtgaagggctacggagcgcgggtcgtcaaggactggctcaactg
gtggggcggctgggccgttggcgagtactgggacaccaacgttgatgcactcctcaactgggcctactcgagcggcgccaaggtcttcgactt
cccgctctactacaagatggatgaggcctttgacaacaaaaacattccagcgctcgtctctgcccttcagaacggccagactgttgtctcccgcg
acccgttcaaggccgtaacctttgtagcaaaccacgacaccgatataatctggaacaagtatccagcctacgcgttcatcctcacctacgaggg
ccagccgacaatattctatcgcgactacgaggagtggctcaacaaggataagctcaagaacctcatctggatacatgacaacctcgccggagg
aagcactgacatcgtttactacgacaacgacgagctgatattcgtgagaaacggctacggaagcaagccgggactgataacatacatcaacct
cgcctcaagcaaagccggaaggtgggtttacgttccgaagttcgcaggctcgtgcatacacgagtacaccggcaatctcggcggctgggtgg

FIGURE 16P

acaagtgggtggactcaagcggctgggtctacctcgaggctcctgcccacgacccggccaacggccagtacggctactccgtctggagctac
tgcggggtggggtga

SEQ ID NO.: 38
Met Ala Lys Tyr Leu Glu Leu Glu Glu Gly Gly Val Ile Met Gln Ala Phe Tyr Trp Asp Val Pro Ser
Gly Gly Ile Trp Trp Asp Thr Ile Arg Gln Lys Ile Pro Glu Trp Tyr Asp Ala Gly Ile Ser Ala Ile Trp
Ile Pro Pro Ala Ser Lys Gly Met Gly Gly Ala Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Phe Phe Asp
Leu Gly Glu Tyr Asp Gln Lys Gly Thr Val Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Val Asn
Met Ile Asn Thr Ala His Ala Tyr Gly Met Lys Val Ile Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly
Asp Leu Glu Trp Asn Pro Phe Val Asn Asp Tyr Thr Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys
Tyr Thr Ala Asn Tyr Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe Gly Gly
Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp Leu Trp Ala Ser Gln Glu Ser Tyr Ala
Ala Tyr Leu Arg Ser Ile Gly Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala Arg Val
Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn Val Asp Ala
Leu Leu Asn Trp Ala Tyr Ser Ser Gly Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Glu
Ala Phe Asp Asn Lys Asn Ile Pro Ala Leu Val Ser Ala Leu Gln Asn Gly Gln Thr Val Val Ser Arg
Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala
Tyr Ala Phe Ile Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu Trp Leu Asn
Lys Asp Lys Leu Lys Asn Leu Ile Trp Ile His Asp Asn Leu Ala Gly Gly Ser Thr Asp Ile Val Tyr
Tyr Asp Asn Asp Glu Leu Ile Phe Val Arg Asn Gly Tyr Gly Ser Lys Pro Gly Leu Ile Thr Tyr Ile
Asn Leu Ala Ser Ser Lys Ala Gly Arg Trp Val Tyr Val Pro Lys Phe Ala Gly Ser Cys Ile His Glu
Tyr Thr Gly Asn Leu Gly Gly Trp Val Asp Lys Trp Val Asp Ser Ser Gly Trp Val Tyr Leu Glu Ala
Pro Ala His Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr Cys Gly Val Gly

SEQ ID NO.: 39
atggccaagtacctggagctcgaagagggcggggtcataatgcaggcgttctactgggacgtgccttcaggaggaatatggtgggacacaat
acggcagaagataccggagtggtacgatgccggaatctccgcaatatggattcctcccgcgagcaggggtatgagcggcggctattcgatgg
gctacgacccctacgattattttgacctcggtgagtactaccagaagggaacggtggaaacgaggttcggctcaaagcaggagctcataaacat
gataaacaccgcccacgcctatggcatgaaggtaatagccgatatagtcatcaaccaccgcgccggcggtgacctggagtggaacccccttcgt
gaacgactatacctggaccgacttctcaaaggtcgcgtcgggtaaatacacggccaactacctcgacttccacccgaacgagctccatgcggg
cgattccggaacatttggaggctatcccgacatatgccacgacaagagctgggaccagtactggctctgggccagccaggagagctacgcgg
catatctcaggagcatcggtatcgatgcctggcgctttgactacgtgaagggctacggagcgtgggtcgtcaaggactggctcaactggtggg
gcggctgggccgttggcgagtactgggaccccaacgttgatgccctcctccccctgggcctactcgagcggcgccaaggtcttcgacttcccgc
tctactacaagatggatgaggcctttgacaacaaaaacattccagcgctcgtctctgccctttcagaacggccagactgttgtctcccgcgacccg
ttcaaggccgtaacctttgtagccaaccacgatacccgatataatctggaacaagtatccagcctacgcgttcatcctcacctacgagggccagcc
gacaatattctaccgcgactacgaggagtcggctcaactaaggataagctcaagaaactcatctggatacaatgacaacctcgccgaggaagca
ccgacatagtctactacgataacgatgaactcatttttgtcaggaacggctacggcagcaagccgggcgttataacctacatcaacctaggctc
gagcaaggccggaaggtgggctacgttccgaagttcgccgggagcgtgcatccacgagtacaccggcaaacctcggcggctgggtggacaa
gtgggtggactcaagcgggtgggtgtacctcgaggccccctgcccacgacccggccaacggctattacggctactccgtctggagctactgcg
gggtgggctga

SEQ ID NO.: 40
Met Ala Lys Tyr Leu Glu Leu Glu Glu Gly Gly Val Ile Met Gln Ala Phe Tyr Trp Asp Val Pro Ser
Gly Gly Ile Trp Trp Asp Thr Ile Arg Gln Lys Ile Pro Glu Trp Tyr Asp Ala Gly Ile Ser Ala Ile Trp
Ile Pro Pro Ala Ser Arg Gly Met Ser Gly Gly Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Tyr Phe Asp
Leu Gly Glu Tyr Tyr Gln Lys Gly Thr Val Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Ile Asn Met
Ile Asn Thr Ala His Ala Tyr Gly Met Lys Val Ile Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly Asp
Leu Glu Trp Asn Pro Phe Val Asn Asp Tyr Thr Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr
Thr Ala Asn Tyr Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe Gly Gly Tyr
Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp Leu Trp Ala Ser Gln Glu Ser Tyr Ala Ala

FIGURE 16Q

Tyr Leu Arg Ser Ile Gly Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp Val Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Pro Asn Val Asp Ala Leu Leu Pro Trp Ala Tyr Ser Ser Gly Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Glu Ala Phe Asp Asn Lys Asn Ile Pro Ala Leu Val Ser Ala Leu Gln Asn Gly Gln Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu Trp Leu Asn Lys Asp Lys Leu Lys Asn Leu Ile Trp Ile His Asp Asn Leu Ala Gly Gly Ser Thr Asp Ile Val Tyr Tyr Asp Asn Asp Glu Leu Ile Phe Val Arg Asn Gly Tyr Gly Asp Lys Pro Gly Leu Ile Thr Tyr Ile Asn Leu Gly Ser Ser Lys Ala Gly Arg Trp Val Tyr Val Pro Lys Phe Ala Gly Ala Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp Val Asp Lys Trp Val Asp Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro Ala His Asp Pro Ala Asn Gly Tyr Tyr Gly Tyr Ser Val Trp Ser Tyr Cys Gly Val Gly

SEQ ID NO.: 41

atggccaagtacctggagctcgaagagggcggggtcataatgcaggcgttctactgggacgtgccttcaggaggaatatggtgggacacaat
acggcagaagataccggagtggtacgatgccggaatctccgcaatatggattcctcccgcgagcaagggtatgagcggcggctattcgatgg
gctacgacccctacgattattttgacctcggtgagtactaccagaagggaacggtggaaacgaggttcggctcaaagcaggagctcataaacat
gataaacacggcccatgcctacggcataaaggtcatagcggacatcgtcataaaccaccgcgcaggcggagacctcgagtggaacccgttc
gttggggactacacctggacggacttctcaaaggtggcctcgggcaaatatactgccaactacctcgacttccacccgaacgagctccatgcg
ggcgattccggaacatttggaggctatcccgacatatgccacgacaagagctgggaccagtactggctctgggccagccaggagagctacgc
ggcatatctcaggagcatcggcatcgatgcctggcgctttgactacgtgaagggctacggagcgtgggtcgtcaaggactggctcaactggtg
gggcggctgggccgttggcgagtactgggacaccaacgttgatgcactcctcaactgggcctactcgagcggcgccaaggtcttcgacttccc
gctctactacaagatggacgcggcctttgacaacaagaacattcccgcactcgtcgaggccctcaagaacgggggcacagtcgtcagccgcg
acccgtttaaggccgtaaccttcgttgcaaaccacgacaccgatataatctggaacaagtatccagcctacgcgttcatcctcacctacgagggc
cagccgacaatattctaccgcgactacgaggagtggctcaacaaggataagctcaagaacctcatctggatacatgacaacctcgccggagg
aagcacgagcatagtttactacgacagcgacgagatgatcttcgtgaggaacggctatggaagcaagcctggccttataacttacatcaacctc
ggctcgagcaaggttggaaggtgggtttatgtgccgaagttcgcgggcgcgtgcatccacgagtatactggtaacctcggaggctgggtagac
aagtacgtctactcaagcggctgggtctatctcgaagctccagcttacgaccctgccaacgggcagtatggctactccgtgtggagctactgcg
gtgttgggtga

SEQ ID NO.: 42

Met Ala Lys Tyr Leu Glu Leu Glu Glu Gly Gly Val Ile Met Gln Ala Phe Tyr Trp Asp Val Pro Ser Gly Gly Ile Trp Trp Asp Thr Ile Arg Gln Lys Ile Pro Glu Trp Tyr Asp Ala Gly Ile Ser Ala Ile Trp Ile Pro Pro Ala Ser Lys Gly Met Ser Gly Gly Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Tyr Phe Asp Leu Gly Glu Tyr Tyr Gln Lys Gly Thr Val Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Ile Asn Met Ile Asn Thr Ala His Ala Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Gly Asp Tyr Thr Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe Gly Gly Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp Leu Trp Ala Ser Gln Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp Val Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn Val Asp Ala Leu Leu Asn Trp Ala Tyr Ser Ser Gly Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Ala Ala Phe Asp Asn Lys Asn Ile Pro Ala Leu Val Glu Ala Leu Lys Asn Gly Gly Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu Trp Leu Asn Lys Asp Lys Leu Lys Asn Leu Ile Trp Ile His Asp Asn Leu Ala Gly Gly Ser Thr Ser Ile Val Tyr Tyr Asp Ser Asp Glu Met Ile Phe Val Arg Asn Gly Tyr Gly Ser Lys Pro Gly Leu Ile Thr Tyr Ile Asn Leu Gly Ser Ser Lys Val Gly Arg Trp Val Tyr Val Pro Lys Phe Ala Gly Ala Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp Val Asp Lys Tyr Val Tyr Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro Ala Tyr Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr Cys Gly Val Gly

FIGURE 16R

SEQ ID NO.: 43

atggccaagtactccgagctggaagagggcggcgttataatgcaggccttctactgggacgtcccaggtggaggaatctggtgggacaccat
caggagcaagataccggagtggtacgaggcgggaatatccgccatttggattcccccggcgagcaagggcatgggcggcgcctattcgatg
ggctacgacccctacgacttctttgacctcggtgagtacgaccagaagggaacggtagagacgcgctttggctccaagcaggagctcgtgaa
catgataaacacggcccatgcctacggcataaaggtcatagcggacatcgtcataaaccaccgcgcaggcggagacctcgagtggaacccg
ttcgttggggactacacctggacggacttctcaaaggtggcctcgggcaaatatactgccaactacctcgacttccaccccaacgaggtcaagt
gctgtgacgagggcacatttggaggcttcccagacatagcccacgagaagagctgggaccagcactggctctgggcgagcgatgagagcta
cgccgcctacctaaggagcatcggcgttgatgcctggcgcttcgactacgtcaagggctacggagcgtgggtcgtcaaggactggctggact
ggtggggaggctgggccgtcggggagtactgggacacaaacgttgatgcactgctcaactgggcctactcgagcgatgcaaaagtcttcgac
ttcccgctctactacaagatggatgaggcctttgacaacaaaaacattccagcgctcgtctctgccttcagaacggccagactgttgtctcccgc
gacccgttcaaggccgtaacctttgtagcaaaccacgacaccgatataatctggaacaagtatccagcctacgcgttcatcctcacctacgagg
gccagccgacaatattctaccgcgactacgaggagtggctcaacaaggataagctcaagaacctcatctggatacatgacaacctcgtcggag
gaagcacgagcatagtttactacgacagcgacgagatgatcttcgtgaggaacggctatggaagcaagcctggccttataacttacatcaacct
cggctcgagcaaggttggaaggtgggtttacgttccgaagttcgcaggctcgtgcatacacgagtacaccggcaatctcggcggctgggtgg
acaagtgggtggactcaagcggctgggtctacctcgaggctcctgcccacgacccggccaacggccagtacggctactccgtctggagctac
tgcggtgttggctga

SEQ ID NO.: 44

Met Ala Lys Tyr Ser Glu Leu Glu Glu Gly Gly Val Ile Met Gln Ala Phe Tyr Trp Asp Val Pro Gly
Gly Gly Ile Trp Trp Asp Thr Ile Arg Ser Lys Ile Pro Glu Trp Tyr Glu Ala Gly Ile Ser Ala Ile Trp
Ile Pro Pro Ala Ser Lys Gly Met Gly Gly Ala Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Phe Phe Asp
Leu Gly Glu Tyr Asp Gln Lys Gly Thr Val Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Val Asn
Met Ile Asn Thr Ala His Ala Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly
Asp Leu Glu Trp Asn Pro Phe Val Gly Asp Tyr Thr Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys
Tyr Thr Ala Asn Tyr Leu Asp Phe His Pro Asn Glu Val Lys Cys Cys Asp Glu Gly Thr Phe Gly
Gly Phe Pro Asp Ile Ala His Glu Lys Ser Trp Asp Gln His Trp Leu Trp Ala Ser Asp Glu Ser Tyr
Ala Ala Tyr Leu Arg Ser Ile Gly Val Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp
Val Val Lys Asp Trp Leu Asp Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn Val Asp
Ala Leu Leu Asn Trp Ala Tyr Ser Ser Asp Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp
Glu Ala Phe Asp Asn Lys Asn Ile Pro Ala Leu Val Ser Ala Leu Gln Asn Gly Gln Thr Val Val Ser
Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro
Ala Tyr Ala Phe Ile Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu Trp Leu
Asn Lys Asp Lys Leu Lys Asn Leu Ile Trp Ile His Asp Asn Leu Val Gly Gly Ser Thr Ser Ile Val
Tyr Tyr Asp Ser Asp Glu Met Ile Phe Val Arg Asn Gly Tyr Gly Ser Lys Pro Gly Leu Ile Thr Tyr
Ile Asn Leu Gly Ser Ser Lys Val Gly Arg Trp Val Tyr Val Pro Lys Phe Ala Gly Ser Cys Ile His
Glu Tyr Thr Gly Asn Leu Gly Gly Trp Val Asp Lys Trp Val Asp Ser Ser Gly Trp Val Tyr Leu Glu
Ala Pro Ala His Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr Cys Gly Val Gly

SEQ ID NO.: 45

atggccaagtactccgacctggaagagggcggcgttataatgcaggccttctactgggacgtcccaggtggaggaatctggtgggacaccat
caggagcaagataccggagtggtacgaggcgggaatatccgccatttggattcccccggcgagcaagggcatgggcggcgcctattcgatg
ggctacgacccctacgacttctttgacctcggtgagtacgaccagaagggaacggtagagacgcgctttggctccaagcaggagctcgtgaa
catgataaacacggcccatgcctacggcataaaggtcatagcggacatcgtcataaaccaccgcgcaggcggagacctcgagtggaacccg
ttcgttggggactacacctggacggacttctcaaaggtggcctcgggcaaatatactgccaactacctcgacttccaccccaacgaggtcaagt
gctgtgacgagggcacatttggaggcttcccagacatagcccacgagaagagctgggaccagcactggctctgggcgagcgatgagagcta
cgccgcctacctaaggagcatcggcgttgatgcctggcgctttgactacgtgaagggctacggagcgtgggtcgtcaaggactggctcaactg
gtggggcggctgggccgttggcgagtactgggacaccaacgttgatgcactcctcaactgggcctactcgagcggcgccaaggtcttcgactt
cccgctctactacaagatggatgaggcctttgacaacaaaaacattccagcgctcgtctctgccttcagaacggccagactgttgtctcccgcg
acccgttcaaggccgtaacctttgtagcaaaccacgacaccgatataatctggaacaagtatccagcctacgcgttcatcctcacctacgaggg
ccagccgacaatattctaccgcgactacgaggagtggctcaacaaggataagctcaagaacctcatctggatacatgacaacctcgccggag

FIGURE 16S

gaagcaccgacatagtctactacgataacgatgaactcatcttcgtcaggaacggctacggggacaagccggggcttataacctacatcaacct
aggctcgagcaaggccggaaggtgggtttatgtgccgaagttcgcgggcgcgtgcatccacgagtatactggtaacctcggaggctgggtag
acaagtacgtctactcaagcggctgggtctatctcgaagctccagcttacgaccctgccaacgggcagtatggctactccgtgtggagctattgc
ggtgttgggtga

SEQ ID NO.: 46
Met Ala Lys Tyr Ser Asp Leu Glu Glu Gly Gly Val Ile Met Gln Ala Phe Tyr Trp Asp Val Pro Gly
Gly Gly Ile Trp Trp Asp Thr Ile Arg Ser Lys Ile Pro Glu Trp Tyr Glu Ala Gly Ile Ser Ala Ile Trp
Ile Pro Pro Ala Ser Lys Gly Met Gly Gly Ala Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Phe Phe Asp
Leu Gly Glu Tyr Asp Gln Lys Gly Thr Val Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Val Asn
Met Ile Asn Thr Ala His Ala Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly
Asp Leu Glu Trp Asn Pro Phe Val Gly Asp Tyr Thr Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys
Tyr Thr Ala Asn Tyr Leu Asp Phe His Pro Asn Glu Val Lys Cys Cys Asp Glu Gly Thr Phe Gly
Gly Phe Pro Asp Ile Ala His Glu Lys Ser Trp Asp Gln His Trp Leu Trp Ala Ser Asp Glu Ser Tyr
Ala Ala Tyr Leu Arg Ser Ile Gly Val Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp
Val Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn Val Asp
Ala Leu Leu Asn Trp Ala Tyr Ser Ser Gly Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp
Glu Ala Phe Asp Asn Lys Asn Ile Pro Ala Leu Val Ser Ala Leu Gln Asn Gly Gln Thr Val Val Ser
Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro
Ala Tyr Ala Phe Ile Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu Trp Leu
Asn Lys Asp Lys Leu Lys Asn Leu Ile Trp Ile His Asp Asn Leu Ala Gly Gly Ser Thr Asp Ile Val
Tyr Tyr Asp Asn Asp Glu Leu Ile Phe Val Arg Asn Gly Tyr Gly Asp Lys Pro Gly Leu Ile Thr Tyr
Ile Asn Leu Gly Ser Ser Lys Ala Gly Arg Trp Val Tyr Val Pro Lys Phe Ala Gly Ala Cys Ile His
Glu Tyr Thr Gly Asn Leu Gly Gly Trp Val Asp Lys Tyr Val Tyr Ser Ser Gly Trp Val Tyr Leu Glu
Ala Pro Ala Tyr Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr Cys Gly Val Gly

SEQ ID NO.: 47
atggccaagtacaccgagctggaagagggcggcgttataatgcaggccttctactgggacgtcccaggtggaggaatctggtgggacaccat
caggagcaagataccggagtggtacgaggcgggaatatccgccatttggattcccccggcgagcaagggcatgggcggcgcctattcgatg
ggctacgacccctacgacttctttgacctcggtgagtacgaccagaagggaacggtagagacgcgctttggctccaagcaggagctcgtgaa
catgataaacaccgcccacgcctatggcatgaaggtaatagccgatatagtcatcaaccaccgcgccggcggtgacctggagtggaacccctt
cgtgaacgactatacctggaccgacttctcaaaggtcgcgtcgggtaaatacacggccaactacctcgacttccaccccaacgaggtcaagtg
ctgtgacgagggcacatttggaggcttcccagacatagcccacgagaagagctgggaccagcactggctctgggcgagcgatgagagctac
gccgcctacctaaggagcatcggcgttgatgcctggcgctttgactacgtgaagggctacggagcgtgggtcgtcaaggactggctcaactgg
tggggcggttgggccgttggcgagtactgggacaccaacgttgatgcactcctcaactgggcctactcgagcggcgccaaggtcttcgacttc
ccgctctactacaagatggatgaggcctttgacaacaaaaacattccagcgctcgtatctgcccttcaggaacggccagactgttgtctccggga
cccgttcaaggccgtaacctttgtagcaaaccacgacaccgatataatctggaacaagtaccttgcttatgctttcatcctcacctacgaaggcca
gcccgtcatattctaccgcgactacgaggagtggctcaacaaggacaggttgaacaacctcatatggatacacgaccacctcgcaggtggaag
cacgagcatagtttactacgacagcgacgagatgatcttcgtgaggaacggctatggaagcaagcctggccttataacttacatcaacctcggct
cgagcaaggttggaaggtgggtttacgttccgaagttcgcaggcccgtgcatacacgagtacaccggcaatctcggcggctgggtggacaag
tgggtggactcaagcggctgggtctacctcgaggctcctgcccacgacccggccaacggccagtacggctactccgtctggagctactgcgg
tgttgggtag

SEQ ID NO.: 48
Met Ala Lys Tyr Thr Glu Leu Glu Glu Gly Gly Val Ile Met Gln Ala Phe Tyr Trp Asp Val Pro Gly
Gly Gly Ile Trp Trp Asp Thr Ile Arg Ser Lys Ile Pro Glu Trp Tyr Glu Ala Gly Ile Ser Ala Ile Trp
Ile Pro Pro Ala Ser Lys Gly Met Gly Gly Ala Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Phe Phe Asp
Leu Gly Glu Tyr Asp Gln Lys Gly Thr Val Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Val Asn
Met Ile Asn Thr Ala His Ala Tyr Gly Met Lys Val Ile Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly
Asp Leu Glu Trp Asn Pro Phe Val Asn Asp Tyr Thr Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys

FIGURE 16T

Tyr Thr Ala Asn Tyr Leu Asp Phe His Pro Asn Glu Val Lys Cys Cys Asp Glu Gly Thr Phe Gly Gly Phe Pro Asp Ile Ala His Glu Lys Ser Trp Asp Gln His Trp Leu Trp Ala Ser Asp Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly Val Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp Val Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn Val Asp Ala Leu Leu Asn Trp Ala Tyr Ser Ser Gly Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Glu Ala Phe Asp Asn Lys Asn Ile Pro Ala Leu Val Ser Ala Leu Gln Asn Gly Gln Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Leu Ala Tyr Ala Phe Ile Leu Thr Tyr Glu Gly Gln Pro Val Ile Phe Tyr Arg Asp Tyr Glu Glu Trp Leu Asn Lys Asp Arg Leu Asn Asn Leu Ile Trp Ile His Asp His Leu Ala Gly Gly Ser Thr Ser Ile Val Tyr Tyr Asp Ser Asp Glu Met Ile Phe Val Arg Asn Gly Tyr Gly Ser Lys Pro Gly Leu Ile Thr Tyr Ile Asn Leu Gly Ser Ser Lys Val Gly Arg Trp Val Tyr Val Pro Lys Phe Ala Gly Pro Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp Val Asp Lys Trp Val Asp Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro Ala His Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr Cys Gly Val Gly

SEQ ID NO.: 53

ATGGCCAAGTACTCCGAGCTGGAAGAGGGCGGCGTTATAATGCAGGCCTTCTACTGGG
ACGTCCCAGGTGGAGGAATCTGGTGGGACACCATCAGGAGCAAGATACCGGAGTGGT
ACGAGGCGGGAATATCCGCCATTTGGATTCCCCCGGCGAGCAAGGGCATGGGCGGCG
CCTATTCGATGGGCTACGACCCCTACGACTTCTTTGACCTCGGTGAGTACGACCAGAA
GGGAACGGTAGAGACGCGCTTTGGCTCCAAGCAGGAGCTCGTGAACATGATAAACAC
GGCCCATGCCTACGGCATAAAGGTCATAGCGGACATCGTCATAAACCACCGCACAGGC
GGAGACCTCGAGTGGAACCCGTTCGTTGGGGACTACACCTGGACGGACTTCTCAAAGG
TGGCCTCGGGCAAATATACTGCCAACTACCTCGACTTCCACCCCAACGAGGTCAAGTG
CTGTGACGAGGGCACATTTGGAGGCTTCCCAGACATAGCCCACGAGAAGAGCTGGGA
CCAGCACTGGCTCTGGGCGAGCGATGAGAGCTACGCCGCCTACCTAAGGAGCATCGGC
GTTGATGCCTGGCGCTTCGACTACGTCAAGGGCTACGGAGCGTGGGTCGTCAAGGACT
GGCTGGACTGGTGGGGAGGCTGGGCCGTCGGGGAGTACTGGGACACAAACGTTGATG
CACTGCTCAACTGGGCCTACTCGAGCGATGCAAAAGTCTTCGACTTCCCGCTCTACTAC
AAGATGGATGAGGCCTTTGACAACAAAAACATTCCAGCGCTCGTCTCTGCCCTTCAGA
ACGGCCAGACTGTTGTCTCCCGCGACCCGTTCAAGGCCGTAACCTTTGTAGCAAACCA
CGACACCGATATAATCTGGAACAAGTATCCAGCCTACGCGTTCATCCTCACCTACGAG
GGCCAGCCGACAATATTCTACCGCGACTACGAGGAGTGGCTCAACAAGGATAAGCTCA
AGAACCTCATCTGGATACATGACAACCTCGCCGGAGGAAGCACTGACATCGTTTACTA
CGACAACGACGAGCTGATATTCGTGAGAAACGGCTACGGAAGCAAGCCGGGACTGAT
AACATACATCAACCTCGCCTCAAGCAAAGCCGGAAGGTGGGTCTACGTTCCGAAGTTC
GCGGGAGCGTGCATCCACGAGTACACCGGCAACCTCGGCGGCTGGGTGGACAAGTGG
GTGGACTCAAGCGGGTGGGTGTACCTCGAGGCCCCTGCCCACGACCCGGCCAACGGCT
ATTACGGCTACTCCGTCTGGAGCTACTGCGGTGTTGGCTGA

SEQ ID NO.: 54

MAKYSELEEGGVIMQAFYWDVPGGGIWWDTIRSKIPEWYEAGISAIWIPPASKGMGGAYS
MGYDPYDFFDLGEYDQKGTVETRFGSKQELVNMINTAHAYGIKVIADIVINHRTGGDLEW
NPFVGDYTWTDFSKVASGKYTANYLDFHPNEVKCCDEGTFGGFPDIAHEKSWDQHWLW
ASDESYAAYLRSIGVDAWRFDYVKGYGAWVVKDWLDWWGGWAVGEYWDTNVDALL
NWAYSSDAKVFDFPLYYKMDEAFDNKNIPALVSALQNGQTVVSRDPFKAVTFVANHDTD
IIWNKYPAYAFILTYEGQPTIFYRDYEEWLNKDKLKNLIWIHDNLAGGSTDIVYYDNDELIF
VRNGYGSKPGLITYINLASSKAGRWVYVPKFAGACIHEYTGNLGGWVDKWVDSSGWVY
LEAPAHDPANGYYGYSVWSYCGVG

SEQ ID NO.: 55

FIGURE 16U

ATGGCCAAGTACCTGGAGCTCGAGGAGGGCGGGGTCATAATGCAGGCGTTCTACTGGG
ACGTGCCTTCAGGAGGAATATGGTGGGACACAATACGGCAGAAGATACCGGAGTGGT
ACGATGCCGGAATCTCCGCAATATGGATTCCCCCGGCGAGCAAGGGCATGGGCGGCGC
CTATTCGATGGGCTACGACCCCTACGACTTCTTTGACCTCGGTGAGTACGACCAGAAG
GGAACGGTAGAGACGCGCTTTGGCTCCAAGCAGGAGCTCGTGAACATGATAAACACC
GCCCACGCCTATGGCATGAAGGTAATAGCCGATATAGTCATCAACCACCGCGCCGGCG
GTGACCTGGAGTGGAACCCCTTCGTGAACGACTATACCTGGACCGACTTCTCAAAGGT
CGCGTCGGGTAAATACACGGCCAACTACCTCGACTTCCACCCGAACGAGCTCCATGCG
GGCGATTCCGGAACATTTGGAGGCTATCCCGACATATGCCACGACAAGAGCTGGGACC
AGTACTGGCTCTGGGCCAGCCAGGAGAGCTACGCGGCATATCTCAGGAGCATCGGCAT
CGATGCCTGGCGCTTTGACTACGTGAAGGGCTACGGAGCGTGGGTCGTCAAGGACTGG
CTCAACTGGTGGGGCGGCTGGGCCGTTGGCGAGTACTGGGACACCAACGTTGATGCAC
TCCTCAACTGGGCCTACTCGAGCGGCGCCAAGGTCTTCGACTTCCCGCTCTACTACAAG
ATGGATGAGGCCTTTGACAACAAAAACATTCCAGCGCTCGTCTCTGCCCTTCAGAACG
GCCAGACTGTTGTCTCCCGCGACCCGTTCAAGGCCGTAACCTTTGTAGCAAACCACGA
CACCGATATAATCTGGAACAAGTACCTTGCTTATGCTTTCATCCTCACCTACGAAGGCC
AGCCCGTCATATTCTACCGCGACTACGAGGAGTGGCTCAACAAGGACAGGTTGAACAA
CCTCATATGGATACACGACCACCTCGCAGGTGGAAGCACGAGCATAGTTTACTACGAC
AGCGACGAGATGATCTTCGTGAGGAACGGCTATGGAAGCAAGCCTGGCCTTATAACTT
ACATCAACCTCGGCTCGAGCAAGGTTGGAAGGTGGGTTTACGTTCCGAAGTTCGCAGG
CTCGTGCATACACGAGTACACCGGCAATCTCGGCGGCTGGGTGGACAAGTGGGTGGAC
TCAAGCGGCTGGGTCTACCTCGAGGCTCCTGCCCACGACCCGGCCAACGGCCAGTACG
GCTACTCCGTCTGGAGCTATTGCGGTGTTGGCTGA

SEQ ID NO.: 56
MAKYLELEEGGVIMQAFYWDVPSGGIWWDTIRQKIPEWYDAGISAIWIPPASKGMGGAYS
MGYDPYDFFDLGEYDQKGTVETRFGSKQELVNMINTAHAYGMKVIADIVINHRAGGDLE
WNPFVNDYTWTDFSKVASGKYTANYLDFHPNELHAGDSGTFGGYPDICHDKSWDQYWL
WASQESYAAYLRSIGIDAWRFDYVKGYGAWVVKDWLNWWGGWAVGEYWDTNVDALL
NWAYSSGAKVFDFPLYYKMDEAFDNKNIPALVSALQNGQTVVSRDPFKAVTFVANHDTD
IIWNKYLAYAFILTYEGQPVIFYRDYEEWLNKDRLNNLIWIHDHLAGGSTSIVYYDSDEMIF
VRNGYGSKPGLITYINLGSSKVGRWVYVPKFAGSCIHEYTGNLGGWVDKWVDSSGWVYL
EAPAHDPANGQYGYSVWSYCGVG

SEQ ID NO.: 57
ATGGCCAAGTACCTGGAGCTCGAAGAGAGCGGGGTCATAATGCAGGCGTTCTACTGGG
ACGTGCCTTCAGGAGGAATATGGTGGGACACAATACGGCAGAAGATACCGGAGTGGT
ACGATGCCGGAATCTCCGCAATATGGATTCCTCCCGCGAGCAAGGGTATGAGCGGCGG
CTATTCGATGGGCTACGACCCCTACGATTATTTGACCTCGGTGAGTACTACCAGAAGG
GAACGGTGGAAACGAGGTTCGGCTCAAAGCAGGAGCTCATAAACATGATAAACACCG
CCCACGCCTACGGCATCAAGGTCATCGCAGACATAGTAATCAACCACCGCGCCGGAGG
AGACCTTGAGTGGAACCCCTTCGTCAATGACTACACCTGGACGGACTTCTCGAAGGTC
GCTTCCGGCAAGTACACGGCCAACTACCTCGACTTCCACCCCAACGAGGTCAAGTGCT
GTGACGAGGGCACATTTGGAGGCTTCCCAGACATAGCCCACGAGAAGAGCTGGGACC
AGCACTGGCTCTGGGCGAGCGATGAGAGCTACGCCGCCTACCTAAGGAGCATCGGCGT
TGATGCCTGGCGCTTTGACTACGTGAAGGGCTACGGAGCGTGGGTCGTCAAGGACTGG
CTCAACTGGTGGGGTGGCTGGGCCGTCGGGGAGTACTGGGACACAAACGTTGATGCAC
TGCTCAACTGGGCCTACTCGAGCGATGCAAAAGTCTTCGACTTCCCGCTCTACTACAAG
ATGGACGAGGCCTTCGATAACAACAACATTCCCGCCCTGGTGGACGCCCTCAGATACG
GTCAGACAGTGGTCAGCCGCGACCCGTTCAAGGCTGTGACGTTTGTAGCCAACCACGA

FIGURE 16V

TACCGATATAATCTGGAACAAGTACCTTGCTTATGCTTTCATCCTCACCTACGAAGGCC
AGCCCGTCATATTCTACCGCGACTACGAGGAGTGGCTCAACAAGGACAGGTTGAACAA
CCTCATATGGATACACGACCACCTCGCAGGTGGAAGCACTGACATCGTTTACTACGAC
AACGACGAGCTGATATTCGTGAGAAACGGCTACGGAAGCAAGCCGGGACTGATAACA
TACATCAACCTCGCCTCAAGCAAAGCCGGAAGGTGGGTCTACGTTCCGAAGTTCGCGG
GAGCGTGCATCCACGAGTACACCGGCAACCTCGGCGGCTGGGTGGACAAGTGGGTGG
ACTCAAGCGGGTGGGTGTACCTCGAGGCCCCTGCCCACGACCCGGCCAACGGCTATTA
CGGCTACTCCGTCTGGAGCTATTGCGGTGTTGGCTGA

SEQ ID NO.: 58
MAKYLELEESGVIMQAFYWDVPSGGIWWDTIRQKIPEWYDAGISAIWIPPASKGMSGGYS
MGYDPYDYFDLGEYYQKGTVETRFGSKQELINMINTAHAYGIKVIADIVINHRAGGDLEW
NPFVNDYTWTDFSKVASGKYTANYLDFHPNEVKCCDEGTFGGFPDIAHEKSWDQHWLW
ASDESYAAYLRSIGVDAWRFDYVKGYGAWVVKDWLNWWGGWAVGEYWDTNVDALL
NWAYSSDAKVFDFPLYYKMDEAFDNNNIPALVDALRYGQTVVSRDPFKAVTFVANHDTD
IIWNKYLAYAFILTYEGQPVIFYRDYEEWLNKDRLNNLIWIHDHLAGGSTDIVYYDNDELIF
VRNGYGSKPGLITYINLASSKAGRWVYVPKFAGACIHEYTGNLGGWVDKWVDSSGWVY
LEAPAHDPANGYYGYSVWSYCGVG

SEQ ID NO.: 59
atggccaagtacctggagctcgaagagggcggggtcataatgcaggcgttctactgggacgtgccttcaggaggaatatggtgggacacaat
acggcagaagataccggagtggtacgatgccggaatctccgcaatatggattcctcccgcgagcaagggtatgagcggcggctattcgatgg
gctacgacccctacgattattttgacctcggtgagtactaccagaagggaacggtggaaacgaggttcggctcaaagcaggagctcataaacat
gataaacaccgcccacgcctacggcatcaaggtcatcgcagacatagtaatcaaccaccgcgccggaggagaccttgagtggaacccctttcg
tcaatgactacacctggacggacttctcgaaggtcgcttccggcaagtacacggccaactacctcgacttcacccgaacgagctccatgcgg
gcgattccggaacatttggaggctatcccgacatatgccacgacaagagctgggaccagtactggctctgggccagccaggagagctacgcg
gcatatctcaggagcatcggcatcgatgcctggcgcttcgactacgtcaagggctatgctccctgggtcgtcaaggactggctgaactggtggg
gaggctgggcggttggagagtactgggacaccaacgtcgacgctgttctcaactgggcatactcgagcggtgccaaggtctttgacttcgccct
ctactacaagatggacgaggccttcgataacaacaacattcccgccctggtggacgccctcagatacggtcagacagtggtcagccgcgacc
cgttcaaggctgtgacgtttgtagccaaccacgatacccgatataatttggaacaagtacccggcctacgccttcatcctcacctacgagggccag
ccgacgatattctaccgcgactacgaggagtggctcaacaaggacaggctcaagaacctcatctggatacacgaccacctcgccggtggaag
cactgacatcgtttactacgacaacgacgagctgatattcgtgagaaacggctacggaagcaagccgggactgataacatacatcaacctcgc
gtcaagcaaagccggaaggtgggtttatgtgccgaagttcgcgggcgcgtgcatccacgagtatactggtaacctcggaggctgggtagaca
agtacgtctactcaagcggctgggtctatctcgaagctccagcttacgaccctgccaacgggcagtatggctactccgtgtggagctattgcggt
gttgggtga

SEQ ID NO.: 60
Met Ala Lys Tyr Leu Glu Leu Glu Glu Gly Gly Val Ile Met Gln Ala Phe Tyr Trp Asp Val Pro Ser
Gly Gly Ile Trp Trp Asp Thr Ile Arg Gln Lys Ile Pro Glu Trp Tyr Asp Ala Gly Ile Ser Ala Ile Trp
Ile Pro Pro Ala Ser Lys Gly Met Ser Gly Gly Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Tyr Phe Asp
Leu Gly Glu Tyr Tyr Gln Lys Gly Thr Val Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Ile Asn Met
Ile Asn Thr Ala His Ala Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly Asp
Leu Glu Trp Asn Pro Phe Val Asn Asp Tyr Thr Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr
Thr Ala Asn Tyr Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe Gly Gly Tyr
Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp Leu Trp Ala Ser Gln Glu Ser Tyr Ala Ala
Tyr Leu Arg Ser Ile Gly Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Ala Pro Trp Val Val
Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn Val Asp Ala Val
Leu Asn Trp Ala Tyr Ser Ser Gly Ala Lys Val Phe Asp Phe Ala Leu Tyr Tyr Lys Met Asp Glu Ala
Phe Asp Asn Asn Asn Ile Pro Ala Leu Val Asp Ala Leu Arg Tyr Gly Gln Thr Val Val Ser Arg Asp
Pro Phe Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr

FIGURE 16W

Ala Phe Ile Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu Trp Leu Asn Lys Asp Arg Leu Lys Asn Leu Ile Trp Ile His Asp His Leu Ala Gly Gly Ser Thr Asp Ile Val Tyr Tyr Asp Asn Asp Glu Leu Ile Phe Val Arg Asn Gly Tyr Gly Ser Lys Pro Gly Leu Ile Thr Tyr Ile Asn Leu Ala Ser Ser Lys Ala Gly Arg Trp Val Tyr Val Pro Lys Phe Ala Gly Ala Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp Val Asp Lys Tyr Val Tyr Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro Ala Tyr Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr Cys Gly Val Gly

SEQ ID NO.: 61

atggccaagtactccgagctgaaaaagggcggggtcataatgcaggcgttctactgggacgtgccttcaggaggaatatggtgggacacaat
acggcagaagataccggagtggtacgaggcgggaatatccgccatttggattcctcccgcgagcaagggtatgagcggcggctattcgatgg
gctacgacccctacgattattttgacctcggtgagtactaccagaagggaacggtggaaacgaggttcggctcaaagcaggagctcataaacat
gataaacaccgcccacgcctacggcatcaaggtcatcgcagacatagtaatcaaccaccgcgccggaggagaccttgagtggaaccccttcg
tcaatgactacacctggacggacttctcgaaggtcgcttccggcaagtacacggccaactacctcaacttccacccgaacgagctccatgcgg
gcgattccggaacatttggaggctatcccgacatatgccacgacaagagctgggaccagtactggctctgggccagccaggagagctacgcg
gcatatctcaggagcatcggcatcgatgcctggcgcttcgactacgtcaagggctacggagcgtgggtcgtcaaggactggctggactggtg
gggaggctgggccgtcggggagtactgggacacaaacgttgatgcactgctcaactgggcctactcgagcgatgcaaaagtcttcgacttccc
gctctactacaagatggatgaggcctttgacaacaaaaacattccagcgctcgtctctgcccttcagaacggccagactgttgtctcccgcgacc
cgttcaaggccgtaacctttgtagcaaaccatgacaccgatataatctggaacaagtatccagcctacgcgttcatcctcacctacgagggccag
ccgacaatattctaccgcgactacgaggagtggctcaacaaggataagctcaagaacctcatctggatacatgacaacctcgccggaggaag
caccgacatagtctactacgataacgatgaactcatcttcgtcaggaacggctacggggacaagccggggcttataacctacatcaacctaggc
tcgagcaaggccggaaggtgggtctacgttccgaagttcgcgggagcgtgcatccacgagtacaccggcaacctcggcggctgggtggaca
agtgggtggactcaagcgggtgggtgtacctcgaggcccctgcccacgacccggccaacggctattacggctactccgtctggagctactgc
ggggtgggctga

SEQ ID NO.: 62

Met Ala Lys Tyr Ser Glu Leu Lys Lys Gly Gly Val Ile Met Gln Ala Phe Tyr Trp Asp Val Pro Ser Gly Gly Ile Trp Trp Asp Thr Ile Arg Gln Lys Ile Pro Glu Trp Tyr Glu Ala Gly Ile Ser Ala Ile Trp Ile Pro Pro Ala Ser Lys Gly Met Ser Gly Gly Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Tyr Phe Asp Leu Gly Glu Tyr Tyr Gln Lys Gly Thr Val Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Ile Asn Met Ile Asn Thr Ala His Ala Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Asn Asp Tyr Thr Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr Leu Asn Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe Gly Gly Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp Leu Trp Ala Ser Gln Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp Val Val Lys Asp Trp Leu Asp Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn Val Asp Ala Leu Leu Asn Trp Ala Tyr Ser Ser Asp Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Glu Ala Phe Asp Asn Lys Asn Ile Pro Ala Leu Val Ser Ala Leu Gln Asn Gly Gln Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu Trp Leu Asn Lys Asp Lys Leu Lys Asn Leu Ile Trp Ile His Asp Asn Leu Ala Gly Gly Ser Thr Asp Ile Val Tyr Tyr Asp Asn Asp Glu Leu Ile Phe Val Arg Asn Gly Tyr Gly Asp Lys Pro Gly Leu Ile Thr Tyr Ile Asn Leu Gly Ser Ser Lys Ala Gly Arg Trp Val Tyr Val Pro Lys Phe Ala Gly Ala Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp Val Asp Lys Trp Val Asp Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro Ala His Asp Pro Ala Asn Gly Tyr Tyr Gly Tyr Ser Val Trp Ser Tyr Cys Gly Val Gly

SEQ ID NO.: 63

atggccaagtacctggagctcgaagagggcggggtcataatgcaggcgttctactgggacgtgccttcaggaggaatatggtgggacacaat
acggcagaagataccggagtggtacgatgccggaatctccgcaatatggattcccccggcgagcaagggcatgggcggcgcctattcgatg
ggctacgacccctacgacttctttgacctcggtgagtacgaccagaagggaacggtagagacgcgcttggctccaagcaggagctcgtgaa
catgataaacacggcccatgcctacggcataaaggccatagcggacatcgtcataaaccaccgcgcaggcggagacctcgagtggaacccg

FIGURE 16X

ttcgttggggactacacctggacggacttctcaaaggtggcctcgggcaaatatactgccaactacctcgacttccaccccaacgaggtcaagt
gctgtgacgagggcacatttggaggcttcccagacatagcccacgagaagagctgggaccagcactggctctgggcgagcgatgagagcta
cgccgcctacctaaggagcatcggcgttgatgcctggcgctttgactacgtgaagggctacggagcgtgggtcgtcaaggactggctcaactg
gtggggcggctgggccgttggcgagtactgggacaccaacgttgatgcactcctcaactgggcctactcgagcggcgccaaggtcttcgactt
cccgctctactacaagatggacgcggcctttgacaacaagaacattcccgcactcgtcgaggccctcaagaacgggggcacagtcgtcagcc
gcgacccgtttaaggccgtaaccttcgttgcaaaccacgacaccgatataatctggaacaagtatccagcctacgcgttcatcctcacctacgag
ggccagccgacaatattctaccgcgactacgaggagtggctcaacaaggataagctcaagaacctcatctggatacatgacaacctcgccgg
aggaagcaccgacatagtctactacgataacgatgaactcatcttcgtcaggaacggctacggggacaagccggggcttataacctacatcaa
cctaggctggagcaaggccggaaggtgggtttatgtgccgaagttcgcgggcgcgtgcatccacgagtatactggtaacctcggaggctggg
tagacaagtacgtctactcaagcggctgggtctatctcgaagctccagcttacgaccctgccaacgggcagtatggctactccgtgtggagcta
ctgcggggtggggtga

SEQ ID NO.: 64
Met Ala Lys Tyr Leu Glu Leu Glu Glu Gly Gly Val Ile Met Gln Ala Phe Tyr Trp Asp Val Pro Ser
Gly Gly Ile Trp Trp Asp Thr Ile Arg Gln Lys Ile Pro Glu Trp Tyr Asp Ala Gly Ile Ser Ala Ile Trp
Ile Pro Pro Ala Ser Lys Gly Met Gly Gly Ala Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Phe Phe Asp
Leu Gly Glu Tyr Asp Gln Lys Gly Thr Val Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Val Asn
Met Ile Asn Thr Ala His Ala Tyr Gly Ile Lys Ala Ile Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly
Asp Leu Glu Trp Asn Pro Phe Val Gly Asp Tyr Thr Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys
Tyr Thr Ala Asn Tyr Leu Asp Phe His Pro Asn Glu Val Lys Cys Cys Asp Glu Gly Thr Phe Gly
Gly Phe Pro Asp Ile Ala His Glu Lys Ser Trp Asp Gln His Trp Leu Trp Ala Ser Asp Glu Ser Tyr
Ala Ala Tyr Leu Arg Ser Ile Gly Val Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp
Val Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn Val Asp
Ala Leu Leu Asn Trp Ala Tyr Ser Ser Gly Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp
Ala Ala Phe Asp Asn Lys Asn Ile Pro Ala Leu Val Glu Ala Leu Lys Asn Gly Gly Thr Val Val Ser
Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro
Ala Tyr Ala Phe Ile Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu Trp Leu
Asn Lys Asp Lys Leu Lys Asn Leu Ile Trp Ile His Asp Asn Leu Ala Gly Gly Ser Thr Asp Ile Val
Tyr Tyr Asp Asn Asp Glu Leu Ile Phe Val Arg Asn Gly Tyr Gly Asp Lys Pro Gly Leu Ile Thr Tyr
Ile Asn Leu Gly Trp Ser Lys Ala Gly Arg Trp Val Tyr Val Pro Lys Phe Ala Gly Ala Cys Ile His
Glu Tyr Thr Gly Asn Leu Gly Gly Trp Val Asp Lys Tyr Val Tyr Ser Ser Gly Trp Val Tyr Leu Glu
Ala Pro Ala Tyr Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr Cys Gly Val Gly

SEQ ID NO.: 65
atggccaagtactccgagctggaagaaggcggcgttataatgcaggccttctactgggacgtcccaggtggaggaatctggtggggcaccat
caggagcaagatacccggagtggtacgaggcgggaatatccgccatttggattcctcccgcgagcaagggtatgagcggcggctattcgatgg
gctacgacccctacgattatttgaccgggtgagtactaccaagaagggtacggtggaaacgaggttcggctcaaagcaggaggctcataaacat
gataaacaccgcccacgctatggcatgaaggtaatagccgatatagtcatcaaccaccgcgccggcggtgacctggagtggaaccccttcgt
gaacgactatacctggaccgacttctcaaaggtcgcgtcgggtaaatacacggccaactacctcgacttccaccccgaacgagctccatgcggg
cgattccggaacatttggaggctatcccgacatatgccacgacaagagctgggaccagtactggctctgggccagccaggagagctacgcgg
catatctcaggagcatcggcatcgatgcctggcgcttcgactacgtcaagggctatgctccctgggtcgtcaaggactggctgaactggtgggg
aggctgggcggttggagagtactgggacaccaacgtcgacgctgttctcaactgggcatactcgagcggtgccaaggtctttgacttcgccctc
tactacaagatggacgaggccttcgataacaacaacattcccgccctggtggacgccctcagatacggtcagacagtggtcagccgcgaccc
gttcaaggctgtgacgtttgtagccaaccacgatacccgatataatttggaacaagtacccggcctacgccttcatcctcacctacgagggccagc
cgacgatattctaccgcgactacgaggagtggctcaacaaggacaggctcaagaacctcatctggatacacgaccacctcgccggtggaagc
acgagcatagtttactacgacagcgacgagatgatcttcgtgaggaacggctatggaagcaagcctggccttataacttacatcaacctcggctc
gagcaaggttggaaggtgggtttacgttccgaagttcgcaggctcgtgcatacacgagtacaccggcaatctcggcggctgggtggacaagt
gggtggactcaagcggctgggtctacctcgaggctcctgcccacgacccggccaacggccagtacggctactccgtctggagctattgcggt
gttggctga

FIGURE 16Y

SEQ ID NO.: 66
Met Ala Lys Tyr Ser Glu Leu Glu Glu Gly Gly Val Ile Met Gln Ala Phe Tyr Trp Asp Val Pro Gly Gly Gly Ile Trp Trp Gly Thr Ile Arg Ser Lys Ile Pro Glu Trp Tyr Glu Ala Gly Ile Ser Ala Ile Trp Ile Pro Pro Ala Ser Lys Gly Met Ser Gly Gly Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Tyr Phe Asp Leu Gly Glu Tyr Tyr Gln Lys Gly Thr Val Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Ile Asn Met Ile Asn Thr Ala His Ala Tyr Gly Met Lys Val Ile Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Asn Asp Tyr Thr Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe Gly Gly Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp Leu Trp Ala Ser Gln Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Ala Pro Trp Val Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn Val Asp Ala Val Leu Asn Trp Ala Tyr Ser Ser Gly Ala Lys Val Phe Asp Phe Ala Leu Tyr Tyr Lys Met Asp Glu Ala Phe Asp Asn Asn Asn Ile Pro Ala Leu Val Asp Ala Leu Arg Tyr Gly Gln Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu Trp Leu Asn Lys Asp Arg Leu Lys Asn Leu Ile Trp Ile His Asp His Leu Ala Gly Gly Ser Thr Ser Ile Val Tyr Tyr Asp Ser Asp Glu Met Ile Phe Val Arg Asn Gly Tyr Gly Ser Lys Pro Gly Leu Ile Thr Tyr Ile Asn Leu Gly Ser Ser Lys Val Gly Arg Trp Val Tyr Val Pro Lys Phe Ala Gly Ser Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp Val Asp Lys Trp Val Asp Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro Ala His Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr Cys Gly Val Gly

SEQ ID NO.: 67
atggccaagtacctggagctcgaagagggcggggtcataatgcaggcgttctactgggacgtgccttcgggaggaatatggtgggacacaat
acggcagaagataccggagtggtacgatgccggaatctccgcaatatggattcctcccgcgagcaagggtatgagcggcggctattcgatgg
gctacgacccctacgattattttgacctcggtgagtactaccagaagggaacggtggaaacgaggttcggctcaaagcaggagctcataaacat
gataaacacggcccatgcctacggcataaaggtcatagcggacatcgtcataaaccaccgcgcaggcggagacctcgagtggaacccgttc
gttggggactacacctggacggacttctcaaaggtggcctcgggcaaatatactgccaactacctcgacttccacccccaacgaggtcaagtgct
gtgacgagggcacatttggaggcttcccagacatagcccacgagaagagctgggaccagcactggctctgggcgagcgatgagagctacg
ccgcctacctaaggagcatcggcgttgatgcctggcgcttcgactacgtcaagggctacggagcgtgggtcgtcaaggactggctggactggt
ggggaggctgggccgtcggggagtactgggacacaaacgttgatgcactgctcaactgggcctactcgagcgatgcaaaagtcttcgacttc
ccgctctactacaagatggacgaggccttcgataacaacaacattcccgccctggtggacgccctcagatacggtcagacagtggtcagccgc
gacccgttcaaggctgtgacgtttgtagccaaccacgataccgatataatctggaacaagtatccagcctacgcgttcatcctcacctacgaggg
ccagccgacaatattctaccgcgactacgaggagtggctcaacaaggataagctcaagaacctcatctggatacatgacaacctcgccggag
gaagcacgagcatagtttactacgacagcgacgagatgatcttcgtgaggaacggctatggaagcaagcctggccttataacttacatcaacct
cggctcgagcaaggttggaaggtgggtctacgttccgaagttcgcgggagcgtgcatccacgagtacaccggcaacctcggcggctgggtg
gacaagtgggtggactcaagcgggtgggtgtacctcgaggccccgtccgacgaccccggccaacggctattacggctactccgtctggagcta
ctgcgtggtgtgggctga

SEQ ID NO.: 68
Met Ala Lys Tyr Leu Glu Leu Glu Glu Gly Gly Val Ile Met Gln Ala Phe Tyr Trp Asp Val Pro Ser Gly Gly Ile Trp Trp Asp Thr Ile Arg Gln Lys Ile Pro Glu Trp Tyr Asp Ala Gly Ile Ser Ala Ile Trp Ile Pro Pro Ala Ser Lys Gly Met Ser Gly Gly Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Tyr Phe Asp Leu Gly Glu Tyr Tyr Gln Lys Gly Thr Val Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Ile Asn Met Ile Asn Thr Ala His Ala Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Gly Asp Tyr Thr Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr Leu Asp Phe His Pro Asn Glu Val Lys Cys Cys Asp Glu Gly Thr Phe Gly Gly Phe Pro Asp Ile Ala His Glu Lys Ser Trp Asp Gln His Trp Leu Trp Ala Ser Asp Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly Val Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp Val Val Lys Asp Trp Leu Asp Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn Val Asp Ala Leu Leu Asn Trp Ala Tyr Ser Ser Asp Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Glu

FIGURE 16Z

Ala Phe Asp Asn Asn Asn Ile Pro Ala Leu Val Asp Ala Leu Arg Tyr Gly Gln Thr Val Val Ser Arg
Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala
Tyr Ala Phe Ile Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu Trp Leu Asn
Lys Asp Lys Leu Lys Asn Leu Ile Trp Ile His Asp Asn Leu Ala Gly Gly Ser Thr Ser Ile Val Tyr
Tyr Asp Ser Asp Glu Met Ile Phe Val Arg Asn Gly Tyr Gly Ser Lys Pro Gly Leu Ile Thr Tyr Ile
Asn Leu Gly Ser Ser Lys Val Gly Arg Trp Val Tyr Val Pro Lys Phe Ala Gly Ala Cys Ile His Glu
Tyr Thr Gly Asn Leu Gly Gly Trp Val Asp Lys Trp Val Asp Ser Ser Gly Trp Val Tyr Leu Glu Ala
Pro Ala His Asp Pro Ala Asn Gly Tyr Tyr Gly Tyr Ser Val Trp Ser Tyr Cys Val Val Gly

SEQ ID NO.: 73
atggctctggaagagggcgggcttataatgcaggccttctactgggacgtcccaggtggaggaatctggtgggacaccatagcccagaagat
acccgactgggcgagcgccgggatttcggcaatatggattcctcccgcgagtaagggcatgagcggcggctattcgatgggctacgaccct
acgatttcttcgacctcggtgagtactaccagaagggaagcgttgagacccgcttcggatcaaaagaggagcttgtgaacatgataaacaccgc
ccatgctcacaacatgaaggtcatagcggacatagtcatcaaccaccgcgccggcggcgacctggagtggaatcctttcaccaacagctacac
ctggaccgatttctcgaaggtcgcgtcgggcaagtacacggccaactacctcgacttccacccgaacgagcttcacgcgggcgattccggaa
catttggaggctatcccgacatatgccacgacaagagctgggaccagcactggctctgggccagcaacgaaagctacgccgcctacctccgg
agcatcggcatcgacgcctggcgcttcgactacgtcaagggctacgctccctgggtcgttaagaactggctgaaccggtggggcggctgggc
ggttggagagtactgggacaccaacgtcgatgcactcctgagctgggcctacgacagcggtgctaaagtcttcgacttcccgctctactacaag
atggacgaggccttcgataacaacaacatccccgccctcgtggacgccctcaagaacggaggcacggtcgtcagccgcgacccgttcaaag
ccgtgaccttcgttgccaaccacgataccaacataatctggaacaagtatccggcctacgccttcatcctcacctatgagggacagccggcaat
attctaccgcgactacgaggagtggctcaacaaggacaggctcaggaacctcatctggatacacgaccacctcgcgggaggaagcacagac
atcatctactacgacagcgacgagcttatcttcgtgagaaacggctacggggacaagccgggactgataacctacatcaacctcggctcaagc
aaggccggaaggtgggtctacgttccgaagttcgcaggctcgtgcatacacgagtacaccggcaacctcggcggctggattgacaagtgggt
tgactcaagcggtcgggtctaccttgaggcccccgcccacgacccggccaacggccagtacggctactccgtatggagctactgcggtgttg
ggtga

SEQ ID NO.: 74
Met Ala Leu Glu Glu Gly Gly Leu Ile Met Gln Ala Phe Tyr Trp Asp Val Pro Gly Gly Gly Ile Trp
Trp Asp Thr Ile Ala Gln Lys Ile Pro Asp Trp Ala Ser Ala Gly Ile Ser Ala Ile Trp Ile Pro Pro Ala
Ser Lys Gly Met Ser Gly Gly Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Phe Phe Asp Leu Gly Glu Tyr
Tyr Gln Lys Gly Ser Val Glu Thr Arg Phe Gly Ser Lys Glu Glu Leu Val Asn Met Ile Asn Thr Ala
His Ala His Asn Met Lys Val Ile Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly Asp Leu Glu Trp
Asn Pro Phe Thr Asn Ser Tyr Thr Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn
Tyr Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe Gly Gly Tyr Pro Asp Ile
Cys His Asp Lys Ser Trp Asp Gln His Trp Leu Trp Ala Ser Asn Glu Ser Tyr Ala Ala Tyr Leu Arg
Ser Ile Gly Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Ala Pro Trp Val Val Lys Asn Trp
Leu Asn Arg Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn Val Asp Ala Leu Leu Ser Trp
Ala Tyr Asp Ser Gly Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Glu Ala Phe Asp
Asn Asn Asn Ile Pro Ala Leu Val Asp Ala Leu Lys Asn Gly Gly Thr Val Val Ser Arg Asp Pro Phe
Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Asn Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile
Leu Thr Tyr Glu Gly Gln Pro Ala Ile Phe Tyr Arg Asp Tyr Glu Glu Trp Leu Asn Lys Asp Arg Leu
Arg Asn Leu Ile Trp Ile His Asp His Leu Ala Gly Gly Ser Thr Asp Ile Ile Tyr Tyr Asp Ser Asp Glu
Leu Ile Phe Val Arg Asn Gly Tyr Gly Asp Lys Pro Gly Leu Ile Thr Tyr Ile Asn Leu Gly Ser Ser
Lys Ala Gly Arg Trp Val Tyr Val Pro Lys Phe Ala Gly Ser Cys Ile His Glu Tyr Thr Gly Asn Leu
Gly Gly Trp Ile Asp Lys Trp Val Asp Ser Ser Gly Arg Val Tyr Leu Glu Ala Pro Ala His Asp Pro
Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr Cys Gly Val Gly

SEQ ID NO.: 75
atggctctggaagagggcgggcttataatgcaggcattctactgggacgtccccatgggaggaatctggtgggacacgatagcccagaagat
acccgactgggcaagcgccgggatttcggcgatatggattcccccgcgagcaagggtatgagcggcggctattcgatgggctacgacccct

FIGURE 16AA

acgattattttgacctcggtgagtactaccagaagggaacggtggaaacaagattcggctcaaagcaggagctcataaacatgataaacaccg
cccacgcctatggcatgaaggtaatagccgatatagtcatcaaccaccgcgccggcggcgatctggagtggaaccccttcgtgaacgactata
cctggaccgacttctcgaaggtcgcgtcgggtaaatacacggccaactacctcgacttccacccgaacgagctccacgcgggcgattccgga
acatttggaggctatcccgacatatgccacgacaagagctgggaccagtactggctctgggccagccaggagagctacgcggcctatctcag
gagcatcggcatcgacgcctggcgcttcgactacgtcaagggctatgctccctgggtcgtcagggactggctgaactggtggggaggctggg
cagttggagagtactgggacaccaacgtcgacgctgttctcaactgggcatactcgagcggtgccaaggtctttgacttcgccctctactacaag
atggacgaggccttcgataacaacaacattcccgccctggtggacgccctcagatacggccagacagtggtcagccgcgacccgttcaaggc
tgtgacgtttgtagccaaccacgatacgacataatctggaacaagtatccagcctacgcgttcatcctcacctacgagggccagccgacaatat
tctaccgcgactacgaggagtggctcaacaaggacaagctcaagaacctcatctggatacatgacaacctcgccggagggagcactgacatc
gtttactacgacaacgacgagctgatattcgtgagaaacggctacggaagcaagccgggactgataacatacatcaacctcggctcaagcaaa
gccggaaggtgggtttacgttccgaagttcgcaggctcgtgcatacacgagtacaccggcaacctcggcggctgggtggacaagtgggtgga
ctcaagcggctgggtttacctcgaggctcctgcccacgacccggccaacggccagtacggctactccgtttggagctattgcggtgttgggtga

SEQ ID NO.: 76
Met Ala Leu Glu Glu Gly Gly Leu Ile Met Gln Ala Phe Tyr Trp Asp Val Pro Met Gly Gly Ile Trp
Trp Asp Thr Ile Ala Gln Lys Ile Pro Asp Trp Ala Ser Ala Gly Ile Ser Ala Ile Trp Ile Pro Pro Ala
Ser Lys Gly Met Ser Gly Gly Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Tyr Phe Asp Leu Gly Glu Tyr
Tyr Gln Lys Gly Thr Val Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Ile Asn Met Ile Asn Thr Ala
His Ala Tyr Gly Met Lys Val Ile Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly Asp Leu Glu Trp
Asn Pro Phe Val Asn Asp Tyr Thr Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn
Tyr Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe Gly Gly Tyr Pro Asp Ile
Cys His Asp Lys Ser Trp Asp Gln Tyr Trp Leu Trp Ala Ser Gln Glu Ser Tyr Ala Ala Tyr Leu Arg
Ser Ile Gly Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Ala Pro Trp Val Val Arg Asp Trp
Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn Val Asp Ala Val Leu Asn Trp
Ala Tyr Ser Ser Gly Ala Lys Val Phe Asp Phe Ala Leu Tyr Tyr Lys Met Asp Glu Ala Phe Asp Asn
Asn Asn Ile Pro Ala Leu Val Asp Ala Leu Arg Tyr Gly Gln Thr Val Val Ser Arg Asp Pro Phe Lys
Ala Val Thr Phe Val Ala Asn His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile
Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu Trp Leu Asn Lys Asp Lys Leu
Lys Asn Leu Ile Trp Ile His Asp Asn Leu Ala Gly Gly Ser Thr Asp Ile Val Tyr Tyr Asp Asn Asp
Glu Leu Ile Phe Val Arg Asn Gly Tyr Gly Ser Lys Pro Gly Leu Ile Thr Tyr Ile Asn Leu Gly Ser Ser
Lys Ala Gly Arg Trp Val Tyr Val Pro Lys Phe Ala Gly Ser Cys Ile His Glu Tyr Thr Gly Asn Leu
Gly Gly Trp Val Asp Lys Trp Val Asp Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro Ala His Asp Pro
Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr Cys Gly Val Gly

SEQ ID NO.: 77
atggctctggaagagggcgggctcataatgcaggccttctactgggacgtccccatgggaggaatctggtgggacacgatagcccagaagat
acccgagtgggcaagcgcgcgggattcggcgatatggatcctccgcgagcaagggtatgagcgggcggctattcgatgggctacgaccct
acgattattttgacctcggtgagtactaccagaagggaacggtggaaacgagggttcggctcaaagcaggagctcataaacatgataaacaccg
cccacgcctatggcatgaaggtaatagccgatatagtcatcaaccaccgcgccggcggtgacctggagtggaaccccttcgtgaacgactata
cctggaccgacttctcaaaggtcgcgtcgggtaaatacacggccaactacctcgacttccacccgaacgagctccatgcgggcgattccggaa
catttggaggctatcccgacatatgccacgacaagagctgggaccagtactggctctgggccagccaggagagctacgcggcatatctcagg
agcatcggcatcgatgcctggcgcttcgactacgtcaagggctatgctccctgggtcgtcaaggactggctgaactggtggggaggctgggc
ggttggagagtactgggacaccaacgtcgacgctgttctcaactgggcatactcgagcggtgccaaggtctttgacttcgccctctactacaaga
tggacgaggccttcgataacaacaacattcccgccctggtggacgccctcagatacggtcagacagtggtcagccgcgacccgttcaaggct
gtgacgtttgtagccaaccacgatacgacataatctggaacaagtatccagcctacgcgttcatcctcacctacgagggccagccgacaatatt
ctaccgcgactacgaggagtggctcaacaaggataagctcaagaacctcatctggatacatgacaacctcgccggagggagcactgacatcg
tttactacgacaacgacgagctgatattcgtgagaaacggctacggaagcaagccgggactgataacatacatcaacctcgcctcaagcaaag
ccggaaggtgggtttacgttccgaagttcgcaggctcgtgcatacacgagtacaccggcaatctcggcggctgggtggacaagtgggtggac
tcaagcggctgggtctacctcgaggctcctgcccacgacccggccaacggccagtacggctactccgtctggagctactgcggtgttgggtga

FIGURE 16BB

SEQ ID NO.: 78

Met Ala Leu Glu Glu Gly Gly Leu Ile Met Gln Ala Phe Tyr Trp Asp Val Pro Met Gly Gly Ile Trp Trp Asp Thr Ile Ala Gln Lys Ile Pro Asp Trp Ala Ser Ala Gly Ile Ser Ala Ile Trp Ile Pro Pro Ala Ser Lys Gly Met Ser Gly Gly Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Tyr Phe Asp Leu Gly Glu Tyr Tyr Gln Lys Gly Thr Val Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Ile Asn Met Ile Asn Thr Ala His Ala Tyr Gly Met Lys Val Ile Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Asn Asp Tyr Thr Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe Gly Gly Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp Leu Trp Ala Ser Gln Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Ala Pro Trp Val Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn Val Asp Ala Val Leu Asn Trp Ala Tyr Ser Ser Gly Ala Lys Val Phe Asp Phe Ala Leu Tyr Tyr Lys Met Asp Glu Ala Phe Asp Asn Asn Asn Ile Pro Ala Leu Val Asp Ala Leu Arg Tyr Gly Gln Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu Trp Leu Asn Lys Asp Lys Leu Lys Asn Leu Ile Trp Ile His Asp Asn Leu Ala Gly Gly Ser Thr Asp Ile Val Tyr Tyr Asp Asn Asp Glu Leu Ile Phe Val Arg Asn Gly Tyr Gly Ser Lys Pro Gly Leu Ile Thr Tyr Ile Asn Leu Ala Ser Ser Lys Ala Gly Arg Trp Val Tyr Val Pro Lys Phe Ala Gly Ser Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp Val Asp Lys Trp Val Asp Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro Ala His Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr Cys Gly Val Gly

SEQ ID NO.: 79

atgaagcctgcgaaactcctcgtctttgtgctcgtagtctctatcctcgcggggctctacgcccagcccgcgggggcggccaagtacctggagc
tcgaagagggcggcgtcataatgcaggcgttctactgggacgtgccttcaggaggaatatggtgggacacaatacggcagaagataccgga
gtggtacgatgccggaatctccgcaatatggattcccccggcgagcaagggcatgggcggcgcctattcgatgggctacgacccctacgactt
ctttgacctcggtgagtacgaccagaagggaacggtagagacgcgctttggctccaagcaggagctcgtgaacatgataaacaccgcccacg
cctacggcatcaaggtcatcgcagacatagtaatcaaccaccgcgccggaggagaccttgagtggaacccccttcgtcaatgactacacctgga
cggacttctcgaaggtcgcttccggcaagtacacggccaactacctcgacttccaccccaacgaggtcaagtgctgcgacgagggcacctttg
gagggttcccggacatagcccacgagaagagctgggaccagtactggctctgggcgagcaacgagagctacgccgcctacctcaggagca
tcggcgttgacgcatggcgcttcgactacgtcaagggctacggagcgtgggtcgtcaaggactggctggactggtggggaggctgggccgt
cggggagtactgggacacaaacgttgatgcactgctcaactgggcctactcgagcgatgcaaaagtcttcgacttcccgctctactacaagatg
gacgcggccttgacaacaagaacattccgcactcgtcgaggccctcaagaacggggcacagtcgtcagccgcgacccgtttaaggccgt
aaccttcgttgcaaaccacgacacggacataatttggaacaagtacccggcctacgccttcatcctcacctacgagggccagccgacgatattc
taccgcgactacgaggagtggctcaacaaggacaggctcaagaacctcatctggatacacgaccacctcgccggtggaagcaccgacatag
tctactacgataacgatgaactcatcttcgtcaggaacggctacggggacaagccggggcttataacctacatcaacctaggctcgagcaagg
ccgggaggtgggtctacgttccgaagttcgcggggagcgtgcatccacgagtacaccggcaacctcggcggctgggtggacaagtgggtgga
ctcaagcgggtgggtgtacctcgaggccctggccacgacccggccaacggctattacggctactccgtctggagctactgcgggtccggct
ga

SEQ ID NO.: 80

Met Lys Pro Ala Lys Leu Leu Val Phe Val Leu Val Val Ser Ile Leu Ala Gly Leu Tyr Ala Gln Pro Ala Gly Ala Ala Lys Tyr Leu Glu Leu Glu Glu Gly Gly Val Ile Met Gln Ala Phe Tyr Trp Asp Val Pro Ser Gly Gly Ile Trp Trp Asp Thr Ile Arg Gln Lys Ile Pro Glu Trp Tyr Asp Ala Gly Ile Ser Ala Ile Trp Ile Pro Pro Ala Ser Lys Gly Met Gly Gly Ala Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Phe Phe Asp Leu Gly Glu Tyr Asp Gln Lys Gly Thr Val Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Val Asn Met Ile Asn Thr Ala His Ala Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Asn Asp Tyr Thr Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr Leu Asp Phe His Pro Asn Glu Val Lys Cys Cys Asp Glu Gly Thr Phe Gly Gly Phe Pro Asp Ile Ala His Glu Lys Ser Trp Asp Gln Tyr Trp Leu Trp Ala Ser Asn Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly Val Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala

FIGURE 16CC

Trp Val Val Lys Asp Trp Leu Asp Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn Val Asp Ala Leu Leu Asn Trp Ala Tyr Ser Ser Asp Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Ala Ala Phe Asp Asn Lys Asn Ile Pro Ala Leu Val Glu Ala Leu Lys Asn Gly Gly Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu Trp Leu Asn Lys Asp Arg Leu Lys Asn Leu Ile Trp Ile His Asp His Leu Ala Gly Gly Ser Thr Asp Ile Val Tyr Tyr Asp Asn Asp Glu Leu Ile Phe Val Arg Asn Gly Tyr Gly Asp Lys Pro Gly Leu Ile Thr Tyr Ile Asn Leu Gly Ser Ser Lys Ala Gly Arg Trp Val Tyr Val Pro Lys Phe Ala Gly Ala Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp Val Asp Lys Trp Val Asp Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro Ala His Asp Pro Ala Asn Gly Tyr Tyr Gly Tyr Ser Val Trp Ser Tyr Cys Gly Val Gly

SEQ ID NO.: 81

atgaagaagtttgtcgccctgttcataaccatgttttcgtagtgagcatggcagtcgttgcacagccagctagcgccgcaaagtattccgagctc
gaagaaggcggcgttataatgcaggccttctactgggacgtcccaggtggaggaatctggtgggacaccatcaggagcaagataccggagt
ggtacgaggcgggaatatccgccatttggattccgccagccagcaaggggatgagcggcggttactcgatgggctacgatccctacgatttctt
tgacctcggcgagtacaaccagaagggaaccatcgaaacgcgctttggctctaaacaggagctcatcaatatgataaacacggcccatgccta
cggcataaaggtcatagcggacatcgtcataaaccaccgcgcaggcggagacctcgagtggaacccgttcgttggggactacacctggacg
gacttctcaaaggtggcctcgggcaaatatactgccaactacctcgacttccaccccaacgaggtcaagtgctgtgacgagggcacatttggag
gcttcccagacatagcccacgagaagagctgggaccagcactggctctgggcgagcgatgagagctacgccgcctacctaaggagcatcg
gcgttgatgcctggcgctttgactacgtgaagggctacggagcgtgggtcgtcaaggactggctcaactggtggggcggctgggccgttggc
gagtactgggacaccaacgttgatgcactcctcaactgggcctactcgagcggcgccaaggtcttcgacttcccgctctactacaagatggatg
aggcctttgacaacaaaaacattccagcgctcgtctctgcccttcagaacggccagactgttgtctcccgcgacccgttcaaggccgtaacctttt
gtagcaaaccacgacaccgatataatctggaacaagtaccttgcttatgctttcatcctcacctacgaaggccagcccgtcatattctaccgcgac
tacgaggagtggctcaacaaggacaggttgaacaacctcatatggatacacgaccacctcgcaggtggaagcacgagcatagtctactacga
cagcgacgagatgatcttcgtgaggaacggctatggaagcaagcctggccttataacttacatcaacctcggctcgagcaaggttggaaggtg
ggtttatgtgccgaagttcgcgggcgcgtgcatccacgagtatactggtaacctcggaggctgggtagacaagtacgtctactcaagcggctg
ggtctatctcgaagctccagcttacgaccctgccaacgggcagtatggctactccgtgtggagctattgcggtgttgggtga

SEQ ID NO.: 82

Met Lys Lys Phe Val Ala Leu Phe Ile Thr Met Phe Phe Val Val Ser Met Ala Val Val Ala Gln Pro Ala Ser Ala Ala Lys Tyr Ser Glu Leu Glu Glu Gly Gly Val Ile Met Gln Ala Phe Tyr Trp Asp Val Pro Gly Gly Gly Ile Trp Trp Asp Thr Ile Arg Ser Lys Ile Pro Glu Trp Tyr Glu Ala Gly Ile Ser Ala Ile Trp Ile Pro Pro Ala Ser Lys Gly Met Ser Gly Gly Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Phe Phe Asp Leu Gly Glu Tyr Asn Gln Lys Gly Thr Ile Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Ile Asn Met Ile Asn Thr Ala His Ala Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Gly Asp Tyr Thr Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr Leu Asp Phe His Pro Asn Glu Val Lys Cys Cys Asp Glu Gly Thr Phe Gly Gly Phe Pro Asp Ile Ala His Glu Lys Ser Trp Asp Gln His Trp Leu Trp Ala Ser Asp Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly Val Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp Val Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn Val Asp Ala Leu Leu Asn Trp Ala Tyr Ser Ser Gly Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Glu Ala Phe Asp Asn Lys Asn Ile Pro Ala Leu Val Ser Ala Leu Gln Asn Gly Gln Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Leu Ala Tyr Ala Phe Ile Leu Thr Tyr Glu Gly Gln Pro Val Ile Phe Tyr Arg Asp Tyr Glu Glu Trp Leu Asn Lys Asp Arg Leu Asn Asn Leu Ile Trp Ile His Asp His Leu Ala Gly Gly Ser Thr Ser Ile Val Tyr Tyr Asp Ser Asp Glu Met Ile Phe Val Arg Asn Gly Tyr Gly Ser Lys Pro Gly Leu Ile Thr Tyr Ile Asn Leu Gly Ser Ser Lys Val Gly Arg Trp Val Tyr Val Pro Lys Phe Ala Gly Ala Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp Val Asp Lys Tyr Val Tyr Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro Ala Tyr Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr Cys Gly Val Gly

FIGURE 16DD

SEQ ID NO.: 83

atggctctggaagacggcgggctcataatgcaggccttctactgggatgttcctggaggaggaatctggtgggacacaatagctcaaaagata
cccgaatgggcaagtgcaggaatctcagcgatatggattccaccagcgagtaagggcatgagcggtggttattccatgggctacgatccctac
gatttctttgacctcggcgagtactatcagaaggggacagttgagacgcgcttcggctcaaaggaagaactggtgaacatgataaacaccgca
cactcctacggcataaaggtgatagcagacatagtcataaaccaccgcgccggtggagaccttgagtggaaccccttcgtgaacgactatacct
ggacagacttctcaaaagtcgcctccggtaaatatacggccaactaccttgacttccacccaaacgagcttcactgttgtgatgaaggtacctttg
gaggataccctgatatatgtcacgacaaaagctgggaccagtactggctctgggcgagcagcgaaagctacgctgcctacctcaggagcata
ggggttgacgcctggcgtttcgactacgtcaagggctacggagcatgggttgttaacgactggctcagctggtggggaggctgggccgttgga
gagtactgggacacgaacgttgatgcactcctcaactgggcatacagcagcggcgccaaggtctttgacttcccgctctactacaagatggacg
aagccttcgacaacaccaacatcccggcattagtggatgcactcagatacggccagacagtggtcagccgcgatcccttcaaggcggtaacttt
cgttgccaaccacgatacagatataatctggaacaagtatccggcttatgcattcatcctacctatgagggacagcctgttatattctaccgcgac
tacgaggagtggctcaacaaggataagcttaacaacctcatctggatacacgatcaccttgctggagggagtactgacattgtttactacgacag
cgacgagcttatctttgtgagaaacggctatggcaccaaaccaggactgataacctatatcaacctcggctcaagcaaagttggaaggtgggtc
tacgttccaaagttcgccggttcatgcatccacgagtacaccggcaacctcggcggttggatagacaagtacgtctcctccagcggctgggtct
atcttgaggccccagcccacgacccggcgaacggctactacggctactccgtatggagctactgcggggttgggtga

SEQ ID NO.: 84

Met Ala Leu Glu Asp Gly Gly Leu Ile Met Gln Ala Phe Tyr Trp Asp Val Pro Gly Gly Gly Ile Trp
Trp Asp Thr Ile Ala Gln Lys Ile Pro Glu Trp Ala Ser Ala Gly Ile Ser Ala Ile Trp Ile Pro Pro Ala
Ser Lys Gly Met Ser Gly Gly Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Phe Phe Asp Leu Gly Glu Tyr
Tyr Gln Lys Gly Thr Val Glu Thr Arg Phe Gly Ser Lys Glu Glu Leu Val Asn Met Ile Asn Thr Ala
His Ser Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly Asp Leu Glu Trp Asn
Pro Phe Val Asn Asp Tyr Thr Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr
Leu Asp Phe His Pro Asn Glu Leu His Cys Cys Asp Glu Gly Thr Phe Gly Gly Tyr Pro Asp Ile Cys
His Asp Lys Ser Trp Asp Gln Tyr Trp Leu Trp Ala Ser Ser Glu Ser Tyr Ala Ala Tyr Leu Arg Ser
Ile Gly Val Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp Val Val Asn Asp Trp Leu
Ser Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn Val Asp Ala Leu Leu Asn Trp Ala
Tyr Ser Ser Gly Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Glu Ala Phe Asp Asn Thr
Asn Ile Pro Ala Leu Val Asp Ala Leu Arg Tyr Gly Gln Thr Val Val Ser Arg Asp Pro Phe Lys Ala
Val Thr Phe Val Ala Asn His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile Leu
Thr Tyr Glu Gly Gln Pro Val Ile Phe Tyr Arg Asp Tyr Glu Glu Trp Leu Asn Lys Asp Lys Leu Asn
Asn Leu Ile Trp Ile His Asp His Leu Ala Gly Gly Ser Thr Asp Ile Val Tyr Tyr Asp Ser Asp Glu
Leu Ile Phe Val Arg Asn Gly Tyr Gly Thr Lys Pro Gly Leu Ile Thr Tyr Ile Asn Leu Gly Ser Ser
Lys Val Gly Arg Trp Val Tyr Val Pro Lys Phe Ala Gly Ser Cys Ile His Glu Tyr Thr Gly Asn Leu
Gly Gly Trp Ile Asp Lys Tyr Val Ser Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro Ala His Asp Pro
Ala Asn Gly Tyr Tyr Gly Tyr Ser Val Trp Ser Tyr Cys Gly Val Gly

SEQ ID NO.: 85

atggctctggaagagggcgggcttataatgcaggcattctattgggacgtcccaggtggaggaatctggtgggacaccatagcccagaagata
cccgaatgggcaagtgcaggaatctcagcgatatggattccaccagcgagtaagggaatgagcggtggttattccatgggctacgatccctac
gatttctttgacctcggcgagtactatcagaaggggacagttgagacgcgcttcggctcaaaggaagaactggtgaacatgataaacaccgca
cactcctacggcataaaggtgatagcggacatagtcataaaccaccgcgccggtggaggcctcgagtggaacccccttcgtgaacgactatacc
tggacagacttctcaaaagtcgcctccggtaaatatacagccaactaccttgacttccacccaaacgagcttcactgttgtgatgaaggtacctttg
gaggataccctgatatatgtcacgacaaaagctgggaccagtactggctctgggcgagcagcgaaagctacgctgcctacctcaggagcata
ggggttgacgcctggtgtttcgactacgtcaagggctacggagcctgggttgttaacgactggctcagctggtggggaggctgggccgttgga
gagtactgggacactaacgttgatgcactcctcaactgggcatacaacagcggcgccaaggtctttgacttcccgctctactacaagatggacg
aagccttcgacaataccaacatccccgctttggtttacgccctcaagaatggcgggacagtggtcagccgcgacccattcaaggcggtaacttt
cgttgccaaccacgatacagatataatctggaacaagtatccggcttatgcattcatccttacctatgagggacagcctgttatattctaccgcgac
tacgaggagtggctcaacaaggataagcttaacaacctcatctggatacacgatcaccttgctggagggagtactgacattgtttactacgacag
cgacgagcttatctttgtgagaaacggctatggcaccaaaccaggactgataacctatatcaacctcggctcaagcaaagctggaaggtgggtc

FIGURE 16EE

tacgttccaaagttcgccggttcatgcatccacgagtacaccggcagcctcggcggttggatagacaagtacgtctcctccagcggctgggtct
accttgaggccccggcccacgacccggccaatggccagtatggctactccgtctggagctattgcggggttgggtga

SEQ ID NO.: 86
Met Ala Leu Glu Glu Gly Gly Leu Ile Met Gln Ala Phe Tyr Trp Asp Val Pro Gly Gly Gly Ile Trp Trp Asp Thr Ile Ala Gln Lys Ile Pro Glu Trp Ala Ser Ala Gly Ile Ser Ala Ile Trp Ile Pro Pro Ala Ser Lys Gly Met Ser Gly Gly Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Phe Phe Asp Leu Gly Glu Tyr Tyr Gln Lys Gly Thr Val Glu Thr Arg Phe Gly Ser Lys Glu Glu Leu Val Asn Met Ile Asn Thr Ala His Ser Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly Gly Leu Glu Trp Asn Pro Phe Val Asn Asp Tyr Thr Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr Leu Asp Phe His Pro Asn Glu Leu His Cys Cys Asp Glu Gly Thr Phe Gly Gly Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp Leu Trp Ala Ser Ser Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly Val Asp Ala Trp Cys Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp Val Val Asn Asp Trp Leu Ser Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn Val Asp Ala Leu Leu Asn Trp Ala Tyr Asn Ser Gly Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Glu Ala Phe Asp Asn Thr Asn Ile Pro Ala Leu Val Tyr Ala Leu Lys Asn Gly Gly Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile Leu Thr Tyr Glu Gly Gln Pro Val Ile Phe Tyr Arg Asp Tyr Glu Glu Trp Leu Asn Lys Asp Lys Leu Asn Asn Leu Ile Trp Ile His Asp His Leu Ala Gly Gly Ser Thr Asp Ile Val Tyr Tyr Asp Ser Asp Glu Leu Ile Phe Val Arg Asn Gly Tyr Gly Thr Lys Pro Gly Leu Ile Thr Tyr Ile Asn Leu Gly Ser Ser Lys Ala Gly Arg Trp Val Tyr Val Pro Lys Phe Ala Gly Ser Cys Ile His Glu Tyr Thr Gly Ser Leu Gly Gly Trp Ile Asp Lys Tyr Val Ser Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro Ala His Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr Cys Gly Val Gly

SEQ ID NO: 87
atgttcctgctcgcgtttttgctcactgcctcgctgttctgcccaacaggacagcccgccaaggctgccgcaccgtttaacggcaccatgatgca
gtattttgaatggtacttgccggatgatggcacgttatggaccaaagtggccaatgaagccaacaacttatccagccttggcatcaccgctctttg
gctgccgcccgcttacaaaggaacaagccgcagcgacgtagggtacggagtatacgacttgtatgacctcggcgaattcaatcaaaaagggga
ccgtccgcacaaaatacggaacaaaagctcaatatcttcaagccattcaagccgcccacgccgctggaatgcaagtgtacgccgatgtcgtgtt
cgaccataaaggcggcgctgacggcacggaatgggtggacgccgtcgaagtcaatccgtccgaccgcaaccaagaaatctcgggcacctat
caaatccaagcatggacgaaatttgattttcccgggcggggcaacacctactccagctttaagtggcgctggtaccattttgacggcgttgattgg
gacgaaagccgaaaattgagccgcatttacaaattccgcggcatcggcaaagcgtgggattgggaagtagacacggaaaacggaaactatg
actacttaatgtatgccgaccttgatatggatcatcccgaagtcgtgaccgagctgaaaaactgggggaaatggtatgtcaacacaacgaacatt
gatgggttccggcttgatgccgtcaagcatattaagttcagtttttttcctgattggttgtcgtatgtgcgttctcagactggcaagccgctatttaccg
tcggggaatattggagctatgacatcaacaagttgcacaattacattacgaaaacagacggaacgatgtctttgtttgatgccccgttacacaaca
aatttatacgcgttccaaatcagggggcgcgatttgatatgcgcacgttaatgacaatatctccatgaaagatcaaccgacatttggccgtcaccctt
cgttgataatcatgacaccgaaccccgccaagcgctgcagtcatgcgtcgaccatgtttcaaacgttggcttacgccttcatcttaatctcggc
aggaaggatacccgtgcgtctttatggtgactattatggcattccacaatataacattcctcgctgaaaagcaaaatcgatccgctcctcatcgc
gcgcagggattatgcttacggaacgcaacatgattatcttgatcactccgacatcatcgggtggacaagggaaggggtcactgaaaaaccagg
atccgggctggccgcactgatcaccgatgggccgggaggaagcaaatggatgtactgttggcaaacaacacgctggaaaagtgttctatga

SEQ ID NO: 88
Met Phe Leu Leu Ala Phe Leu Leu Thr Ala Ser Leu Phe Cys Pro Thr Gly Gln Pro Ala Lys Ala Ala Ala Pro Phe Asn Gly Thr Met Met Gln Tyr Phe Glu Trp Tyr Leu Pro Asp Asp Gly Thr Leu Trp Thr Lys Val Ala Asn Glu Ala Asn Asn Leu Ser Ser Leu Gly Ile Thr Ala Leu Trp Leu Pro Pro Ala Tyr Lys Gly Thr Ser Arg Ser Asp Val Gly Tyr Gly Val Tyr Asp Leu Tyr Asp Leu Gly Glu Phe Asn Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly Thr Lys Ala Gln Tyr Leu Gln Ala Ile Gln Ala Ala His Ala Ala Gly Met Gln Val Tyr Ala Asp Val Val Phe Asp His Lys Gly Gly Ala Asp Gly Thr Glu Trp Val Asp Ala Val Glu Val Asn Pro Ser Asp Arg Asn Gln Glu Ile Ser Gly Thr Tyr Gln Ile Gln Ala Trp Thr Lys Phe Asp Phe Pro Gly Arg Gly Asn Thr Tyr Ser Ser Phe Lys Trp Arg Trp Tyr His Phe Asp

FIGURE 16FF

Gly Val Asp Trp Asp Glu Ser Arg Lys Leu Ser Arg Ile Tyr Lys Phe Arg Gly Ile Gly Lys Ala Trp Asp Trp Glu Val Asp Thr Glu Asn Gly Asn Tyr Asp Tyr Leu Met Tyr Ala Asp Leu Asp Met Asp His Pro Glu Val Val Thr Glu Leu Lys Asn Trp Gly Lys Trp Tyr Val Asn Thr Thr Asn Ile Asp Gly Phe Arg Leu Asp Ala Val Lys His Ile Lys Phe Ser Phe Phe Pro Asp Trp Leu Ser Tyr Val Arg Ser Gln Thr Gly Lys Pro Leu Phe Thr Val Gly Glu Tyr Trp Ser Tyr Asp Ile Asn Lys Leu His Asn Tyr Ile Thr Lys Thr Asp Gly Thr Met Ser Leu Phe Asp Ala Pro Leu His Asn Lys Phe Tyr Thr Ala Ser Lys Ser Gly Gly Ala Phe Asp Met Arg Thr Leu Met Thr Asn Thr Leu Met Lys Asp Gln Pro Thr Leu Ala Val Thr Phe Val Asp Asn His Asp Thr Glu Pro Gly Gln Ala Leu Gln Ser Trp Val Asp Pro Trp Phe Lys Pro Leu Ala Tyr Ala Phe Ile Leu Thr Arg Gln Glu Gly Tyr Pro Cys Val Phe Tyr Gly Asp Tyr Tyr Gly Ile Pro Gln Tyr Asn Ile Pro Ser Leu Lys Ser Lys Ile Asp Pro Leu Leu Ile Ala Arg Arg Asp Tyr Ala Tyr Gly Thr Gln His Asp Tyr Leu Asp His Ser Asp Ile Ile Gly Trp Thr Arg Glu Gly Val Thr Glu Lys Pro Gly Ser Gly Leu Ala Ala Leu Ile Thr Asp Gly Pro Gly Gly Ser Lys Trp Met Tyr Cys Trp Gln Thr Thr Arg Trp Lys Ser Val Leu

SEQ ID NO: 89

atgaaagaagcggttgtgtatcaaattttcccggatcggttctttaatggcaacccttcaaatgataacagcaagcagcaggcacgcggggcgc
agccgattgagcatcgcgattggtcggatttgcccgataatccgcgcctgaaagggacgagcggctacgatggcgacggtgaatggtcgaat
gacttttcggcggagacatcgccggaattgaacaaaagttggattatttgcagtcgcttggagtgaacacgatttacttaaatccgatcgccaatg
cgccatcgaaccataaatatgatgcgagcaattacaaagaattggatccgatgttcggttcccggaagaattccaatcgtttgtgcaggcgcttg
cgaaccgggggatgcatctcatcttagacggggtgttcaaccacgtatccgacgattcgatttactttgaccgctaccaccgctatccgaccgtc
ggtgcgtatgaatattgggaagcggtttacgatttgatgaatgaaaaaggattgagcgaggaagaagcgcggaaacaagtggaagagaagttc
aaacaagagggacagacgttcagcccgtatgggtttcatcttggttcaatattgaaaacaaaaaagtcaatggccattatcaataccaatcatggt
ggggctatgacagtctgccggagtttaagtcggtgacgggggaaaaagtgccgcatccgagtgaattgaacaacgatgcgctcgcgaattac
attttccgtgaatcggattcggtggcgaaaagctggattgccctcggcgcctccggctggcggttggatgtggccaatgaggtggatccggcgt
ttggcgcgagtttcgccaagaattgcttcaagggtcgtacggccgcggtccgacgttaaaagaggggggagcagccgctcatttttaggggaaa
tttgggatgacgcatcgaaatattttctaggcgaccagtacgattccgtgatgaactaccggttccgcggggcggtgcttgacttttgaaaaacg
gaaatgcagaagaggcggacaagcggctgacggccataagggaagactacccaagtgaagcgttttatgcgctgatgaacttaatcggttcg
catgacacggcgcgggcggtctttctgcttgggaacggaacggattcatccgagcgggcggagcttgatccgaattataatgaggaacttggg
aaaaagcggctcaagctggcggtgattttgcagatgggatacccgggagcgccgacgatttattacggcgatgaagcgggagtaacaggctc
aaaagacccagacaaccgccgcacgtatccgtggggcaaagaagatcaaaatctgttgtcccattatcagaaagtggggcacattcgccagc
accatcaatcgttgttggcccatggcgacatcaagacggtgtatgcgcaaggggatgtatacgtatttgcccgccaatacgggcgtgaagcgg
cgctcattgccatcaaccgcggcaatgaggacaagacggtggcgcttgacgtcgcttcgttgcttccgaacggcaccgtgcttacggatgagtt
gcatgatggcggggaagctacggtcgctggcggaacgttgacggtcacgattccggccctggatggacggatgatgtttgggacggtgacg
gcggaaatgccggcagcagtcagcaatttgcaggcgagcgcttcggatggctgcgtgacgttaacgtgggaaggaaatgcatcgagatacc
gaatttacgagtccacgttaaaaggtgccggttatacgatggtgcaagagacggaaacaacttcggccacgatcggttcgttgacgaacggaa
cagcctattacttggcgttgcggcggtcgatgaaaagggaatgaatcaccgaaggttcgaaacgaatcgcgtcgttcctcattacccgcttgac
gagcgacaatgtccagttcgtaacaacgttagcgatgaccaactggattgtcaagccgcaagaaggtgatccatgtaacatgacaat
gtgacaagcaaaggagcagctgatgggttgcaagcggtgttgcaagtgaaaggccgcatgacgaaacatggaaagaatacagagcggctt
accaaggacaagacggcgacgccaacgtgttccgagctgccttcactccgctcgccgcagggacgtatacgtatcgttatgcgctgacgacc
aaccttggcgaggagtggatgtatacagaagagaagcaagtgacgtttgcggcagacaacagcgaccaaatagcgccagcagacgccatcg
agctgcggcagcctgcggttgaatcgggacaagtgaatttatcatggacgtttgttgggaaaaaagatggggatgcttatttgttagccatcgag
cgcaacggtgatatcgtgcatacaaccacttcgatcggcgattcatttacagactacgatgtcgaaaacggcaccgagtacacgtatgttgtcaa
gttgtatgaccgcgccggcaatgttgtggcgtcaaacacggtcaaggtgacgccggacattgtgatggtgaaagtgatttttaaagtgagagcg
ccggattacacaccgttggatgcccgaattacgattccgaacagcttgaacggctggaacacaggggcctgggagatgtcgcgcaacggtgc
ggtgacgcccgattggcaatttaccgtcgaggtgcaggaaggggaaacgatcacctataagtatgtgaaaggcggatcgtgggatcaagagg
ggttggccgaccatacgcgtgaggacgacaacgatgatgacgtgagctactacggctatgggacgattggcaccgacttgaaagtgacggtc
cacaatgaaggaaacaatacgatgattgtgcaagacccgcattttgcgctggatcgatatgccggtcgtcatcgaagaggtgcaaaaacaagga
agtcaagtgacgatcaagggcaatgccattaaaaacggtgttttgacgatcaatggcgagcgggtgccgattgatggccggatggcattctcgt
acacgtttgcgccggccagccatcaaaaagaagtgttgatccatatcgaaccatcggccgaaagcaaaacagccatttttcaacaacgacggcg

FIGURE 16GG

gagcgattgcgaaaaacacaaaagattacgtgctgaatttagaaacgaagcaattcaaaaagcttctcgagagtacttctagagcggccgcgg
gcccatcgattttccacccgggtggggtaccaggta

SEQ ID NO: 90
Met Lys Glu Ala Val Val Tyr Gln Ile Phe Pro Asp Arg Phe Phe Asn Gly Asn Pro Ser Asn Asp Asn Ser Lys Gln Gln Ala Arg Gly Ala Gln Pro Ile Glu His Arg Asp Trp Ser Asp Leu Pro Asp Asn Pro Arg Leu Lys Gly Thr Ser Gly Tyr Asp Gly Asp Gly Glu Trp Ser Asn Asp Phe Phe Gly Gly Asp Ile Ala Gly Ile Glu Gln Lys Leu Asp Tyr Leu Gln Ser Leu Gly Val Asn Thr Ile Tyr Leu Asn Pro Ile Ala Asn Ala Pro Ser Asn His Lys Tyr Asp Ala Ser Asn Tyr Lys Glu Leu Asp Pro Met Phe Gly Ser Pro Glu Glu Phe Gln Ser Phe Val Gln Ala Leu Ala Asn Arg Gly Met His Leu Ile Leu Asp Gly Val Phe Asn His Val Ser Asp Asp Ser Ile Tyr Phe Asp Arg Tyr His Arg Tyr Pro Thr Val Gly Ala Tyr Glu Tyr Trp Glu Ala Val Tyr Asp Leu Met Asn Glu Lys Gly Leu Ser Glu Glu Glu Ala Arg Lys Gln Val Glu Glu Lys Phe Lys Gln Glu Gly Gln Thr Phe Ser Pro Tyr Gly Phe His Leu Trp Phe Asn Ile Glu Asn Lys Lys Val Asn Gly His Tyr Gln Tyr Gln Ser Trp Trp Gly Tyr Asp Ser Leu Pro Glu Phe Lys Ser Val Thr Gly Glu Lys Val Pro His Pro Ser Glu Leu Asn Asn Asp Ala Leu Ala Asn Tyr Ile Phe Arg Glu Ser Asp Ser Val Ala Lys Ser Trp Ile Ala Leu Gly Ala Ser Gly Trp Arg Leu Asp Val Ala Asn Glu Val Asp Pro Ala Phe Trp Arg Glu Phe Arg Gln Glu Leu Leu Gln Gly Ser Tyr Gly Arg Gly Pro Thr Leu Lys Glu Gly Glu Gln Pro Leu Ile Leu Gly Glu Ile Trp Asp Asp Ala Ser Lys Tyr Phe Leu Gly Asp Gln Tyr Asp Ser Val Met Asn Tyr Arg Phe Arg Gly Ala Val Leu Asp Phe Leu Lys Asn Gly Asn Ala Glu Glu Ala Asp Lys Arg Leu Thr Ala Ile Arg Glu Asp Tyr Pro Ser Glu Ala Phe Tyr Ala Leu Met Asn Leu Ile Gly Ser His Asp Thr Ala Arg Ala Val Phe Leu Leu Gly Asn Gly Thr Asp Ser Ser Glu Arg Ala Glu Leu Asp Pro Asn Tyr Asn Glu Glu Leu Gly Lys Lys Arg Leu Lys Leu Ala Val Ile Leu Gln Met Gly Tyr Pro Gly Ala Pro Thr Ile Tyr Tyr Gly Asp Glu Ala Gly Val Thr Gly Ser Lys Asp Pro Asp Asn Arg Arg Thr Tyr Pro Trp Gly Lys Glu Asp Gln Asn Leu Leu Ser His Tyr Gln Lys Val Gly His Ile Arg Gln His His Gln Ser Leu Leu Ala His Gly Asp Ile Lys Thr Val Tyr Ala Gln Gly Asp Val Tyr Val Phe Ala Arg Gln Tyr Gly Arg Glu Ala Ala Leu Ile Ala Ile Asn Arg Gly Asn Glu Asp Lys Thr Val Ala Leu Asp Val Ala Ser Leu Leu Pro Asn Gly Thr Val Leu Thr Asp Glu Leu His Asp Gly Gly Glu Ala Thr Val Ala Gly Gly Thr Leu Thr Val Thr Ile Pro Ala Leu Asp Gly Arg Met Met Phe Gly Thr Val Thr Ala Glu Met Pro Ala Ala Val Ser Asn Leu Gln Ala Ser Ala Ser Asp Gly Cys Val Thr Leu Thr Trp Glu Gly Asn Ala Ser Arg Tyr Arg Ile Tyr Glu Ser Thr Leu Lys Gly Ala Gly Tyr Thr Met Val Gln Glu Thr Glu Thr Thr Ser Ala Thr Ile Gly Ser Leu Thr Asn Gly Thr Ala Tyr Tyr Phe Ala Val Ala Ala Val Asp Glu Asn Gly Asn Glu Ser Pro Lys Val Glu Thr Asn Arg Val Val Pro His Tyr Pro Leu Thr Ser Asp Asn Val Gln Phe Val Thr Thr Leu Ser Asp Ala Thr Leu Asp Leu Ser Lys Pro Gln Gln Val Asp Val His Val Asn Ile Asp Asn Val Thr Ser Lys Gly Ala Ala Asp Gly Leu Gln Ala Val Leu Gln Val Lys Gly Pro His Asp Glu Thr Trp Lys Glu Tyr Arg Ala Ala Tyr Gln Gly Gln Asp Gly Asp Ala Asn Val Phe Arg Ala Ala Phe Thr Pro Leu Ala Ala Gly Thr Tyr Thr Tyr Arg Tyr Ala Leu Thr Thr Asn Leu Gly Glu Glu Trp Met Tyr Thr Glu Glu Lys Gln Val Thr Phe Ala Ala Asp Asn Ser Asp Gln Ile Ala Pro Ala Asp Ala Ile Glu Leu Arg Gln Pro Ala Val Glu Ser Gly Gln Val Asn Leu Ser Trp Thr Phe Val Gly Lys Lys Asp Gly Asp Ala Tyr Leu Leu Ala Ile Glu Arg Asn Gly Asp Ile Val His Thr Thr Thr Ser Ile Gly Asp Ser Phe Thr Asp Tyr Asp Val Glu Asn Gly Thr Glu Tyr Thr Tyr Val Val Lys Leu Tyr Asp Arg Ala Gly Asn Val Val Ala Ser Asn Thr Val Lys Val Thr Pro Asp Ile Val Met Val Lys Val Ile Phe Lys Val Arg Ala Pro Asp Tyr Thr Pro Leu Asp Ala Arg Ile Thr Ile Pro Asn Ser Leu Asn Gly Trp Asn Thr Gly Ala Trp Glu Met Ser Arg Asn Gly Ala Val Thr Pro Asp Trp Gln Phe Thr Val Glu Val Gln Glu Gly Glu Thr Ile Thr Tyr Lys Tyr Val Lys Gly Gly Ser Trp Asp Gln Glu Gly Leu Ala Asp His Thr Arg Glu Asp Asp Asn Asp Asp Asp Val Ser Tyr Tyr Gly Tyr Gly Thr Ile Gly Thr Asp Leu Lys Val Thr Val His Asn Glu Gly Asn Asn Thr Met Ile Val Gln Asp Arg Ile Leu Arg Trp Ile Asp Met Pro Val Val Ile Glu Glu Val Gln Lys Gln Gly Ser Gln Val Thr Ile Lys Gly Asn Ala Ile Lys Asn Gly Val Leu Thr Ile Asn Gly Glu Arg Val Pro Ile Asp Gly Arg Met Ala Phe Ser Tyr Thr Phe Ala Pro Ala Ser His Gln Lys Glu Val Leu Ile His Ile Glu Pro Ser Ala Glu Ser Lys Thr Ala Ile Phe Asn Asn Asp

FIGURE 16HH

Gly Gly Ala Ile Ala Lys Asn Thr Lys Asp Tyr Val Leu Asn Leu Glu Thr Lys Gln Phe Lys Lys Leu Leu Glu Ser Thr Ser Arg Ala Ala Ala Gly Pro Ser Ile Phe His Pro Gly Gly Val Pro Gly

SEQ ID NO: 91

gtgctaacgtttcaccgcatcattcgaaaaggatggatgttcctgctcgcgtttttgctcactgcctcgctgttctgcccaacaggacagcccgcca
aggctgccgcaccgtttaacggcaccatgatgcagtattttgaatggtacttgccggatgatggcacgttatggaccaaagtggccaatgaagc
caacaacttatccagccttggcatcaccgctctttggctgccgcccgcttataaaggaacaagccgcagcgacgtagggtacggagtatacga
cttgtatgacctcggcgaattcaatcaaaaagggaccgtccgcacaaaatacggaacaaaagctcaatatcttcaagccattcaagccgcccac
gccgctggaatgcaagtgtacgccgatgtcgtgttcgaccataaggcggcgccgacggcacggaatgggtggacgccgtcgaagtcaatc
cgtccgaccgcaaccaagaaatctcgggcacctatcaaatccaagcatggacgaaatttgattttccgggcggggcaacacctactccagctt
taagtggcgctggtaccattttgacggcgttgattgggacgaaagccgaaaattgagccgcatttacaaattccgcggcatcggcaaagcgtgg
gattgggaagtagacacggaaaacggaaactatgactacttaatgtatgccgacttggacatggaccatcctgaagtggttacggaactgaaaa
actggggcaaatggtatgtcaacacaacgaacattgatgggttccggcttgatgccgtcaagcatattaagttcagttttttcctgattggttgtcgt
atgtgcgttctcagactggcaagccgctatttaccgtcggggaatattggagctatgacatcaacaagttgcacaattacattacgaaaacaaacg
gaacgatgtctttgtttgatgccccgttacacaacaaattttataccgcttccaaatcaggggggcgcatttgatatgcgcacgttaatgaccaatact
ctcatgaaagatcaaccgacattggccgtcaccttcgttgataatcatgacaccgaacccggccaagcgctgcagtcatgggtcgacccatggt
tcaaaccgttggcttacgcctttattctaactcggcaggaaggatacccgtgcgtcttttatggtgactattatggcatcccacaatataacattcctt
cgctgaaaagcaaaatcgatccgctcctcatcgcgcgcagggattatgcttacggaacgcaacatgattatcttgatcactccgacatcatcggg
tggacaagggaaggcgtcactgaaaaaccaggatccggactggccgcactgatcaccgatgggccgggaggaagcaaatggatgtacgtt
ggcaaacaacacgccggaaaagtgttctatgaccttaccggcaaccggagtgacaccgtcaccatcaacagtgatggatggggagaattcaa
agtcaatggcggttcggttcggtttgggttcctagaaaaacgaccgtctctaccatcgcttggccgatcacaacccgaccgtggactggtgaatt
cgtccgttggaccgaaccacggttggtggcatggccttga

SEQ ID NO: 92

Val Leu Thr Phe His Arg Ile Ile Arg Lys Gly Trp Met Phe Leu Leu Ala Phe Leu Leu Thr Ala Ser Leu Phe Cys Pro Thr Gly Gln Pro Ala Lys Ala Ala Ala Pro Phe Asn Gly Thr Met Met Gln Tyr Phe Glu Trp Tyr Leu Pro Asp Asp Gly Thr Leu Trp Thr Lys Val Ala Asn Glu Ala Asn Asn Leu Ser Ser Leu Gly Ile Thr Ala Leu Trp Leu Pro Pro Ala Tyr Lys Gly Thr Ser Arg Ser Asp Val Gly Tyr Gly Val Tyr Asp Leu Tyr Asp Leu Gly Glu Phe Asn Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly Thr Lys Ala Gln Tyr Leu Gln Ala Ile Gln Ala Ala His Ala Ala Gly Met Gln Val Tyr Ala Asp Val Val Phe Asp His Lys Gly Gly Ala Asp Gly Thr Glu Trp Val Asp Ala Val Glu Val Asn Pro Ser Asp Arg Asn Gln Glu Ile Ser Gly Thr Tyr Gln Ile Gln Ala Trp Thr Lys Phe Asp Phe Pro Gly Arg Gly Asn Thr Tyr Ser Ser Phe Lys Trp Arg Trp Tyr His Phe Asp Gly Val Asp Trp Asp Glu Ser Arg Lys Leu Ser Arg Ile Tyr Lys Phe Arg Gly Ile Gly Lys Ala Trp Asp Trp Glu Val Asp Thr Glu Asn Gly Asn Tyr Asp Tyr Leu Met Tyr Ala Asp Leu Asp Met Asp His Pro Glu Val Val Thr Glu Leu Lys Asn Trp Gly Lys Trp Tyr Val Asn Thr Thr Asn Ile Asp Gly Phe Arg Leu Asp Ala Val Lys His Ile Lys Phe Ser Phe Phe Pro Asp Trp Leu Ser Tyr Val Arg Ser Gln Thr Gly Lys Pro Leu Phe Thr Val Gly Glu Tyr Trp Ser Tyr Asp Ile Asn Lys Leu His Asn Tyr Ile Thr Lys Thr Asn Gly Thr Met Ser Leu Phe Asp Ala Pro Leu His Asn Lys Phe Tyr Thr Ala Ser Lys Ser Gly Gly Ala Phe Asp Met Arg Thr Leu Met Thr Asn Thr Leu Met Lys Asp Gln Pro Thr Leu Ala Val Thr Phe Val Asp Asn His Asp Thr Glu Pro Gly Gln Ala Leu Gln Ser Trp Val Asp Pro Trp Phe Lys Pro Leu Ala Tyr Ala Phe Ile Leu Thr Arg Gln Glu Gly Tyr Pro Cys Val Phe Tyr Gly Asp Tyr Tyr Gly Ile Pro Gln Tyr Asn Ile Pro Ser Leu Lys Ser Lys Ile Asp Pro Leu Leu Ile Ala Arg Arg Asp Tyr Ala Tyr Gly Thr Gln His Asp Tyr Leu Asp His Ser Asp Ile Ile Gly Trp Thr Arg Glu Gly Val Thr Glu Lys Pro Gly Ser Gly Leu Ala Ala Leu Ile Thr Asp Gly Pro Gly Gly Ser Lys Trp Met Tyr Val Gly Lys Gln His Ala Gly Lys Val Phe Tyr Asp Leu Thr Gly Asn Arg Ser Asp Thr Val Thr Ile Asn Ser Asp Gly Trp Gly Glu Phe Lys Val Asn Gly Gly Ser Val Ser Val Trp Val Pro Arg Lys Thr Thr Val Ser Thr Ile Ala Trp Pro Ile Thr Thr Arg Pro Trp Thr Gly Glu Phe Val Arg Trp Thr Glu Pro Arg Leu Val Ala Trp Pro

FIGURE 16II

SEQ ID NO: 93

atgaaatcgtttgcattcatgcctatcctttttatgcaaacgatttcatcagtgaaagggaaggaggaggaaaaatggggaagaatatgagaaga
agattcacgtatttttcaatcttcttattgttcgttcagctgttttcatttagtgcaaccgctagcgccaatggaacggtgaacagtagtcctgtggttaa
tggaaacgaagtcacgtttctatatggaggaacaggaaacgagcagtctgtgttactggcaggctcctttaatgattggcagaaagatggtgaca
agaagattgcactaacaaaaggcgacaataacgtctggtctgtcacgcaaacacttcaagatgggacatatacgtataagtttgttgtagatggtc
aatgggtggcggatccgcttaacccgaatcaagtagacgacggttacggcggccgtaatagtgtcgttgttgtcgggacaccggtgcaacaag
aacggacagtgacgcttgttggtaacttacaagacgaattaggtcatacgagcgaatgggatccgaaagcgacagctacagtgatgaaaaagg
aagggaacgggttatatacgtttacaggtacacttccagccggaacgtacgagtataaaattgcgattaatggcagctgggacgaaaactatggt
gtcggcggccgcgatggcgggaatattaagctgctattaaatgaacaaacaacggttacattttattacaacgacagaacgcatgcgattgcgg
attcgacttggtatgcaccaattctaaaagaaaagcagccgcggctcgttggaacgattttaccagctattggttatgaaacagacgtgaacggtt
ggacgccgcaaacatcaacggcgttgttgtcagatgatgattttgattccatttatacgtttaaggcgcgtgtgccaaaagggacatatgaatataa
agtagttcttgggaatgattggacatatgaaaattatccacaagataatgccaaattaaatgtgcttgaagaaacgacaattacctttttctttaacgc
gaaaacgaaagtagtgtataccgattacaatccaagcggttcggatggtatcgtccaaaaagaccgtttgaagcataatacgtgggattcgttgta
tcgccaaccgtttggtgcggtgaaagctgggacagaagtgacccttcgtttatcagcgaaaaaaggtgatttgacaaaagcggatgtatatgtaa
aaaatacgacaaccggcacagcgaaactatattcgatgaaaaaagccggtgttcttggcgaagaagaatattgggaagcgacattcacaccgg
atgtgaaaggagtatacggttataaatttattgcggtagatgctggaacgaaagcagaatacggggaagatacacaagaagggcagtgggga
aaagcagtagataaaaatgcagagctgttccaattaacggtgtacgacccatcctaccaaacaccggattggatgaaagaagcagttgtatatca
aattttccctgatccaaag

SEQ ID NO: 94

Met Lys Ser Phe Ala Phe Met Pro Ile Leu Phe Tyr Ala Asn Asp Phe Ile Ser Glu Arg Glu Gly Gly
Gly Lys Met Gly Lys Asn Met Arg Arg Arg Phe Thr Tyr Phe Ser Ile Phe Leu Leu Phe Val Gln Leu
Phe Ser Phe Ser Ala Thr Ala Ser Ala Asn Gly Thr Val Asn Ser Ser Pro Val Val Asn Gly Asn Glu
Val Thr Phe Leu Tyr Gly Gly Thr Gly Asn Glu Gln Ser Val Leu Leu Ala Gly Ser Phe Asn Asp Trp
Gln Lys Asp Gly Asp Lys Lys Ile Ala Leu Thr Lys Gly Asp Asn Asn Val Trp Ser Val Thr Gln Thr
Leu Gln Asp Gly Thr Tyr Thr Tyr Lys Phe Val Val Asp Gly Gln Trp Val Ala Asp Pro Leu Asn Pro
Asn Gln Val Asp Asp Gly Tyr Gly Gly Arg Asn Ser Val Val Val Val Gly Thr Pro Val Gln Gln Glu
Arg Thr Val Thr Leu Val Gly Asn Leu Gln Asp Glu Leu Gly His Thr Ser Glu Trp Asp Pro Lys Ala
Thr Ala Thr Val Met Lys Lys Glu Gly Asn Gly Leu Tyr Thr Phe Thr Gly Thr Leu Pro Ala Gly Thr
Tyr Glu Tyr Lys Ile Ala Ile Asn Gly Ser Trp Asp Glu Asn Tyr Gly Val Gly Gly Arg Asp Gly Gly
Asn Ile Lys Leu Leu Leu Asn Glu Gln Thr Thr Val Thr Phe Tyr Tyr Asn Asp Arg Thr His Ala Ile
Ala Asp Ser Thr Trp Tyr Ala Pro Ile Leu Lys Glu Lys Gln Pro Arg Leu Val Gly Thr Ile Leu Pro
Ala Ile Gly Tyr Glu Thr Asp Val Asn Gly Trp Thr Pro Gln Thr Ser Thr Ala Leu Leu Ser Asp Asp
Asp Phe Asp Ser Ile Tyr Thr Phe Lys Ala Arg Val Pro Lys Gly Thr Tyr Glu Tyr Lys Val Val Leu
Gly Asn Asp Trp Thr Tyr Glu Asn Tyr Pro Gln Asp Asn Ala Lys Leu Asn Val Leu Glu Glu Thr
Thr Ile Thr Phe Phe Phe Asn Ala Lys Thr Lys Val Val Tyr Thr Asp Tyr Asn Pro Ser Gly Ser Asp
Gly Ile Val Gln Lys Asp Arg Leu Lys His Asn Thr Trp Asp Ser Leu Tyr Arg Gln Pro Phe Glu Ala
Val Lys Ala Gly Thr Glu Val Thr Leu Arg Leu Ser Ala Lys Lys Gly Asp Leu Thr Lys Ala Asp Val
Tyr Val Lys Asn Thr Thr Thr Gly Thr Ala Lys Leu Tyr Ser Met Lys Lys Ala Gly Val Leu Gly Glu
Glu Glu Tyr Trp Glu Ala Thr Phe Thr Pro Asp Val Lys Gly Val Tyr Gly Tyr Lys Phe Ile Ala Val
Asp Ala Gly Thr Lys Ala Glu Tyr Gly Glu Asp Thr Gln Glu Gly Gln Trp Gly Lys Ala Val Asp Lys
Asn Ala Glu Leu Phe Gln Leu Thr Val Tyr Asp Pro Ser Tyr Gln Thr Pro Asp Trp Met Lys Glu Ala
Val Val Tyr Gln Ile Phe Pro Asp Pro Lys

SEQ ID NO: 95

atgtatacactattcatccgttcatattttgatactgatggtgatggtgtaggagactttagtggagttgctgaaaaggtagattatctaaaatctcttg
gagtagatacagtctggtttttaccatttaataaaagtaaatcttatcatggatatgatgttgaagattactatgatgtagaaccagattatggaacact
acaagatcttgataatatgataaaagttctaaatgaaaatggaataaaggtagtaatggatcttgttgttaatcatacgtcggatacacatccatggtt
tcttgatgcagttgaaaatactactaattctccatattggaactattacattatgagcttggatgagcctcaaaataagaatcattggcattataaggtt
aattcaaaaggacaaactgtgtggtatttggattgtttgattcatcaatgccggaccttaattacgacaaccctaaagtaatggatgaagtgaaaaa

FIGURE 16JJ

aataatagatttttgggcagatatgggagtagatggatttagattagatgcagcaaaacattattatggatttgactggagcgatggaattgaacag
tcagcaagcgttgcaaaagagatagaagactatataaaagataaactaggggaaaatgcaatagttgtgagtgaggtttacgatggagattcaa
atgttcttttaaaatttgctccaatgcctgtgtttaattttagttttatgtacaatttgagaggaaatttgaagggagagataacttaatttcagactctatt
agttgggttgattcctcgttgtataatttaaatgtttttcatttccatttattgatagtcatgatcttgacagatttatttctgagcttgtagatagtaaatatc
agggagatgtaatatctgccacaaaacaatatttgctagttaatgctttactactctcattaacaggcatgccaactatttactatggtgatgaaatag
gacttaggggatggaagtggcattcagaaccatgggatatacctgtgcgtgagccaatgcaatggtataaggatcaaaaagggaacggtcaaa
cttattggacaaaagagttttacgaaggtattactgaaggaagtgctaatgaagatggagcaatatacgatgatccagatgatggagtatctgtag
aagaacaagaaaatggatattctatttttaaactttttaaagaatttatcaacttacgaaaagattatccggcacttgcttttggaagtactacgattga
gagagattggaaaaacttgtatgtttgaaaaagtcgtataacttccaggatgttcttgtattaattaaccttgatccaacgtattcaaatacatacgaa
gttccagaagggtataaatgggtgtggtatgcatttttgatggtgacaactatgaatttggagcaaaagatgaaatgattttacagaatacaagttg
gacgataaatccaaggcaaatttatatatttgtaaagtaa

SEQ ID NO: 96
Met Tyr Thr Leu Phe Ile Arg Ser Tyr Phe Asp Thr Asp Gly Asp Gly Val Gly Asp Phe Ser Gly Val
Ala Glu Lys Val Asp Tyr Leu Lys Ser Leu Gly Val Asp Thr Val Trp Phe Leu Pro Phe Asn Lys Ser
Lys Ser Tyr His Gly Tyr Asp Val Glu Asp Tyr Tyr Asp Val Glu Pro Asp Tyr Gly Thr Leu Gln Asp
Leu Asp Asn Met Ile Lys Val Leu Asn Glu Asn Gly Ile Lys Val Val Met Asp Leu Val Val Asn His
Thr Ser Asp Thr His Pro Trp Phe Leu Asp Ala Val Glu Asn Thr Thr Asn Ser Pro Tyr Trp Asn Tyr
Tyr Ile Met Ser Leu Asp Glu Pro Gln Asn Lys Asn His Trp His Tyr Lys Val Asn Ser Lys Gly Gln
Thr Val Trp Tyr Phe Gly Leu Phe Asp Ser Ser Met Pro Asp Leu Asn Tyr Asp Asn Pro Lys Val Met
Asp Glu Val Lys Lys Ile Ile Asp Phe Trp Ala Asp Met Gly Val Asp Gly Phe Arg Leu Asp Ala Ala
Lys His Tyr Tyr Gly Phe Asp Trp Ser Asp Gly Ile Glu Gln Ser Ala Ser Val Ala Lys Glu Ile Glu
Asp Tyr Ile Lys Asp Lys Leu Gly Glu Asn Ala Ile Val Val Ser Glu Val Tyr Asp Gly Asp Ser Asn
Val Leu Leu Lys Phe Ala Pro Met Pro Val Phe Asn Phe Ser Phe Met Tyr Asn Leu Arg Gly Asn
Phe Glu Gly Arg Asp Asn Leu Ile Ser Asp Ser Ile Ser Trp Val Asp Ser Ser Leu Tyr Asn Leu Asn
Val Phe His Phe Pro Phe Ile Asp Ser His Asp Leu Asp Arg Phe Ile Ser Glu Leu Val Asp Ser Lys
Tyr Gln Gly Asp Val Ile Ser Ala Thr Lys Gln Tyr Leu Leu Val Asn Ala Leu Leu Leu Ser Leu Thr
Gly Met Pro Thr Ile Tyr Tyr Gly Asp Glu Ile Gly Leu Arg Gly Trp Lys Trp His Ser Glu Pro Trp
Asp Ile Pro Val Arg Glu Pro Met Gln Trp Tyr Lys Asp Gln Lys Gly Asn Gly Gln Thr Tyr Trp Thr
Lys Glu Phe Tyr Glu Gly Ile Thr Glu Gly Ser Ala Asn Glu Asp Gly Ala Ile Tyr Asp Asp Pro Asp
Asp Gly Val Ser Val Glu Glu Gln Glu Asn Gly Tyr Ser Ile Leu Asn Phe Phe Lys Glu Phe Ile Asn
Leu Arg Lys Asp Tyr Pro Ala Leu Ala Phe Gly Ser Thr Thr Ile Glu Arg Asp Trp Lys Asn Leu Tyr
Val Leu Lys Lys Ser Tyr Asn Phe Gln Asp Val Leu Val Leu Ile Asn Leu Asp Pro Thr Tyr Ser Asn
Thr Tyr Glu Val Pro Glu Gly Tyr Lys Trp Val Trp Tyr Ala Phe Phe Asp Gly Asp Asn Tyr Glu Phe
Gly Ala Lys Asp Glu Met Ile Leu Gln Asn Thr Ser Trp Thr Ile Asn Pro Arg Gln Ile Tyr Ile Phe Val
Lys

SEQ ID NO: 97
atgaggaagaagatgtcgcattcaagatttacttttcttttgatcttagcacttttattttcttctccggttgtatttcagaagttaaaagcgaaagccag
ctactaaattcaaagcaaaaggtccttgtaaaagtaaatgttaatacgccatttattgagaatgctactactaatacgtggagtgtttcaaaagaatct
tttattgattatcttagtaaagtgattattactgttaaggatgtaaatgatcagattgtatttactaaggaaacaacgaacaaaacaaatatttatttgaa
attgaacttcttcctggaacttatacatttgaggtaaaaggatatgaggaagatttagttatattttcaggggaaaaagttaatcagatcatagatgag
aaaaataatattgttaatgtcgaaacttttttgttaatggaatagttaggacaataattgaagttgacgatattatttataaaaattatgatattacatcgg
caacgttgatcttcaaaaaagatacagcacaagaagattatgaagaggtacctgtaacacttacaggtacttccacttttaattaataaagaattatat
cctggtatgtggactgtaaaatttgaagttgatcttaaatcaaaggatgcaagtatgttaccagaaaaagttcatcttgaaaatgaatttagcataga
agtgcttccagcaaagacaaaaagtttaacatttaatgtagtctttgatacagaggttaatgaaccgaaattagtagttgtatttccgcaaattgagtt
gccttttgtggatcctgtaacaaatttaagtggagagataaatgaattagaagggaatctttcaatgaattgggactattcagatccaaatgcagaat
ttatgtgtataaagaattagaggaacaaggagaatatttgtatgaatttgttggaaaaacacgcgagaaaagttatacaatagaaaattttaccaag
caagaattcgataaatttagtggaatcgctattaatgtttatgccaacggtaaagagagtggattagttgttctaaaaaaagaaaatattaaacttatc
gatttagaaagtgttgacagtataagtgctacttataacgttgatacgaatgagcttaagttggattggaattataccaattcaagtgttactttgaag

FIGURE 16KK

ttttgaaaaaaggtataaatagcaatgaatacgaaataatttctcaactaacacaaaattcttttcaacagaattcacaggcaggcaattttgggatc
ttgagaaaattgcgattagagtagttgctaatggatttgaaagtaagattaatgagatttcaagagatgatataactataacatcattgaatcttcctct
tacatcgtctactatgtatacactattcatccgttcatattttgatactgatggtgatggtgtaggagactttagtggagttgctgaaaaggtagattatc
taaaatctcttggagtagatacagtctggtttttaccatttaataaaagtaaatcttatcatggatatgatgttgaagattactatgatgtagaaccagat
tatggaacactacaagatcttgataatatgataaaagttctaaatgaaaatggaataaaggtagtaatggatcttgttgttaatcatacgtcggatac
acatccatggtttcttgatgcagttgaaaatactactaattctccatattggaactattacattatgagcttggatgagcctcaaaataagaatcattgg
cattataaggttaattcaaaaggacaaactgtgtggtattttggattgtttgattcatcaatgccggaccttaattacgacaaccctaaagtaatggat
gaagtgaaaaaaataatagatttttgggcagatatgggagtagatggatttagattagatgcagcaaaacattattatggatttgactggagcgatg
gaattgaacagtcagcaagcgttgcaaaagagatagaagactatataaaagataaactaggggaaaatgcaatagttgtgagtgaggtttacga
tggagattcaaatgttcttttaaaatttgctccaatgcctgtgtttaattttagttttatgtacaatttgagaggaaatttgaagggagagataacttaatt
tcagactctattagttgggttgattcctcgttgtataatttaaatgtttttcattttccatttattgatagtcatgatcttgacagatttatttctgagcttgtag
atagtaaatatcagggagatgtaatatctgccacaaaacaatatttgctagttaatgctttactactctcattaacaggcatgccaactatttactatgg
tgatgaaataggacttaggggatggaagtggcattcagaaccatgggatataccigtgcgtgagccaatgcaatggtataaggatcaaaaagg
gaacggtcaaacttattggacaaaagagttttacgaaggtattactgaaggaagtgctaatgaagatggagcaatatacgatgatccagatgatg
gagtatctgtagaagaacaagaaaatggatattctatttttaaacttttttaaagaatttatcaacttacgaaaagattatccggcacttgcttttggaagt
actacgattgagagagattggaaaaacttgtatgttttgaaaaagtcgtataacttccaggatgttcttgtattaattaaccttgatccaacgtattcaa
atacatacgaagttccagaagggtataaatgggtgtggtatgcatttttgatggtgacaactatgaatttggagcaaaagatgaaatgattttacag
aatacaagttggacgataaatccaaggcaaatttatatatttgtaaagtaa

SEQ ID NO: 98

Met Arg Lys Lys Met Ser His Ser Arg Phe Thr Phe Leu Leu Ile Leu Ala Leu Phe Ile Phe Phe Ser
Gly Cys Ile Ser Glu Val Lys Ser Glu Ser Gln Leu Leu Asn Ser Lys Gln Lys Val Leu Val Lys Val
Asn Val Asn Thr Pro Phe Ile Glu Asn Ala Thr Thr Asn Thr Trp Ser Val Ser Lys Glu Ser Phe Ile
Asp Tyr Leu Ser Lys Val Ile Ile Thr Val Lys Asp Val Asn Asp Gln Ile Val Phe Thr Lys Glu Thr
Thr Asn Lys Thr Asn Ile Tyr Phe Glu Ile Glu Leu Leu Pro Gly Thr Tyr Thr Phe Glu Val Lys Gly
Tyr Glu Glu Asp Leu Val Ile Phe Ser Gly Glu Lys Val Asn Gln Ile Ile Asp Glu Lys Asn Asn Ile
Val Asn Val Glu Thr Phe Phe Val Asn Gly Ile Val Arg Thr Ile Ile Glu Val Asp Asp Ile Ile Tyr Lys
Asn Tyr Asp Ile Thr Ser Ala Thr Leu Ile Phe Lys Lys Asp Thr Ala Gln Glu Asp Tyr Glu Glu Val
Pro Val Thr Leu Thr Gly Thr Ser Thr Leu Ile Asn Lys Glu Leu Tyr Pro Gly Met Trp Thr Val Lys
Phe Glu Val Asp Leu Lys Ser Lys Asp Ala Ser Met Leu Pro Glu Lys Val His Leu Glu Asn Glu Phe
Ser Ile Glu Val Leu Pro Ala Lys Thr Lys Ser Leu Thr Phe Asn Val Val Phe Asp Thr Glu Val Asn
Glu Pro Lys Leu Val Val Val Phe Pro Gln Ile Glu Leu Pro Phe Val Asp Pro Val Thr Asn Leu Ser
Gly Glu Ile Asn Glu Leu Glu Gly Asn Leu Ser Met Asn Trp Asp Tyr Ser Asp Pro Asn Ala Glu Phe
Tyr Val Tyr Lys Glu Leu Glu Glu Gln Gly Glu Tyr Leu Tyr Glu Phe Val Gly Lys Thr Arg Glu Lys
Ser Tyr Thr Ile Glu Asn Phe Thr Lys Gln Glu Phe Asp Lys Phe Ser Gly Ile Ala Ile Asn Val Tyr
Ala Asn Gly Lys Glu Ser Gly Leu Val Val Leu Lys Lys Glu Asn Ile Lys Leu Ile Asp Leu Glu Ser
Val Asp Ser Ile Ser Ala Thr Tyr Asn Val Asp Thr Asn Glu Leu Lys Leu Asp Trp Asn Tyr Thr Asn
Ser Ser Val Thr Phe Glu Val Leu Lys Lys Gly Ile Asn Ser Asn Glu Tyr Glu Ile Ile Ser Gln Leu Thr
Gln Asn Ser Phe Ser Thr Glu Phe Thr Gly Arg Gln Phe Trp Asp Leu Glu Lys Ile Ala Ile Arg Val
Val Ala Asn Gly Phe Glu Ser Lys Ile Asn Glu Ile Ser Arg Asp Asp Ile Thr Ile Thr Ser Leu Asn Leu
Pro Leu Thr Ser Ser Thr Met Tyr Thr Leu Phe Ile Arg Ser Tyr Phe Asp Thr Asp Gly Asp Gly Val
Gly Asp Phe Ser Gly Val Ala Glu Lys Val Asp Tyr Leu Lys Ser Leu Gly Val Asp Thr Val Trp Phe
Leu Pro Phe Asn Lys Ser Lys Ser Tyr His Gly Tyr Asp Val Glu Asp Tyr Tyr Asp Val Glu Pro Asp
Tyr Gly Thr Leu Gln Asp Leu Asp Asn Met Ile Lys Val Leu Asn Glu Asn Gly Ile Lys Val Val Met
Asp Leu Val Val Asn His Thr Ser Asp Thr His Pro Trp Phe Leu Asp Ala Val Glu Asn Thr Thr Asn
Ser Pro Tyr Trp Asn Tyr Tyr Ile Met Ser Leu Asp Glu Pro Gln Asn Lys Asn His Trp His Tyr Lys
Val Asn Ser Lys Gly Gln Thr Val Trp Tyr Phe Gly Leu Phe Asp Ser Ser Met Pro Asp Leu Asn Tyr
Asp Asn Pro Lys Val Met Asp Glu Val Lys Lys Ile Ile Asp Phe Trp Ala Asp Met Gly Val Asp Gly
Phe Arg Leu Asp Ala Ala Lys His Tyr Tyr Gly Phe Asp Trp Ser Asp Gly Ile Glu Gln Ser Ala Ser
Val Ala Lys Glu Ile Glu Asp Tyr Ile Lys Asp Lys Leu Gly Glu Asn Ala Ile Val Val Ser Glu Val

FIGURE 16LL

Tyr Asp Gly Asp Ser Asn Val Leu Leu Lys Phe Ala Pro Met Pro Val Phe Asn Phe Ser Phe Met Tyr Asn Leu Arg Gly Asn Phe Glu Gly Arg Asp Asn Leu Ile Ser Asp Ser Ile Ser Trp Val Asp Ser Ser Leu Tyr Asn Leu Asn Val Phe His Phe Pro Phe Ile Asp Ser His Asp Leu Asp Arg Phe Ile Ser Glu Leu Val Asp Ser Lys Tyr Gln Gly Asp Val Ile Ser Ala Thr Lys Gln Tyr Leu Leu Val Asn Ala Leu Leu Leu Ser Leu Thr Gly Met Pro Thr Ile Tyr Tyr Gly Asp Glu Ile Gly Leu Arg Gly Trp Lys Trp His Ser Glu Pro Trp Asp Ile Pro Val Arg Glu Pro Met Gln Trp Tyr Lys Asp Gln Lys Gly Asn Gly Gln Thr Tyr Trp Thr Lys Glu Phe Tyr Glu Gly Ile Thr Glu Gly Ser Ala Asn Glu Asp Gly Ala Ile Tyr Asp Asp Pro Asp Asp Gly Val Ser Val Glu Glu Gln Glu Asn Gly Tyr Ser Ile Leu Asn Phe Phe Lys Glu Phe Ile Asn Leu Arg Lys Asp Tyr Pro Ala Leu Ala Phe Gly Ser Thr Thr Ile Glu Arg Asp Trp Lys Asn Leu Tyr Val Leu Lys Lys Ser Tyr Asn Phe Gln Asp Val Leu Val Leu Ile Asn Leu Asp Pro Thr Tyr Ser Asn Thr Tyr Glu Val Pro Glu Gly Tyr Lys Trp Val Trp Tyr Ala Phe Phe Asp Gly Asp Asn Tyr Glu Phe Gly Ala Lys Asp Glu Met Ile Leu Gln Asn Thr Ser Trp Thr Ile Asn Pro Arg Gln Ile Tyr Ile Phe Val Lys

SEQ ID NO: 99

atgtacacactcttcatccgctcttttacgatacaaacaacgacggtgtaggtgactacaacggtgttgcccaaaaagtagactatctcaaaacg
cttggagtggatacagtttggttcttgccgttcaacaaagcaaaatcgtaccacggttacgatgttgaagactactacgatgtagaacctgactatg
gaacatacgcacaacttgaaaatatgataaagacactcaatcagaacggaattcgtgttgttatggacttggttgtgaaccacacttccgatacac
actcgtggtttctggatgccgttgagaacacaacgaattcgaaatattggagctactacataatgacacttgaaaatagagacggttggaatcact
ggcattggaagataaactcaaaagggcaaaaagtttactacttcggactgtttgactcatcaatgcccgatttgaatttcgacaatccacaagtgat
gaacgaaatcaagagaataatcgatttctggataacagttggtgtggatggtttcagacttgatgcaccaaagcactacaaaggctgggattggg
acgacggcatttcaggttcagcagcaatcgcgagggaaatagaaagttacatcaggagcaagttaggaaacgatgcgatagttgtcggggaa
gtgtacgatggaaatccatcggttctttcacaatttgcaccgatgccggcgttcaacttcacattcatgtatggaataacaggcaaccatgagggg
aaagataacctgctgggagaaacaatttcatgggttaatggagcgagttattatctcaacgtaaaacatttcccgttcatagacaatcacgatttga
acagatggatatcgatacttatcgaccaaaagtatagtggaaacacacaagttggtacgaagcagtatattttaacaaatgcgctcttgctttcctta
aacggtatgcctgttatttattatgggaatgaaataggcttgagaggatggaaatggggacaagacccgtgggatttgccggtgagagagccga
tgcagtggtacgcaagtcaaagtggagctgggcagacatggtggacaaagcctgtctaccagcaaaaggaatcacatttggaaatgcaaac
gtcgatggtgcgatgtacgatgatccaaatgatggggtttcagtagaagagcagatgaatggttacacgataaataacttctttaaacaattcataa
ccctgaggaagacatatccggctctatcgaaaggttcgataacgatagaacgcgactggaagaacctgtacgttatcaaacgagtctacggaa
atcaggaagtgcttgtattgataaacttagacccaacttggccgaacaattacacgttaccaggtggatacaggtgggtctggtatgcgttctttaa
tgggagtttgtttgaatttggcaataaaaacgaatcaccactgagccaagataccaactggacagtcaatccaaggcaagtgtatgtgtttgtgaa
ggactaa

SEQ ID NO: 100

Met Tyr Thr Leu Phe Ile Arg Ser Phe Tyr Asp Thr Asn Asn Asp Gly Val Gly Asp Tyr Asn Gly Val Ala Gln Lys Val Asp Tyr Leu Lys Thr Leu Gly Val Asp Thr Val Trp Phe Leu Pro Phe Asn Lys Ala Lys Ser Tyr His Gly Tyr Asp Val Glu Asp Tyr Tyr Asp Val Glu Pro Asp Tyr Gly Thr Tyr Ala Gln Leu Glu Asn Met Ile Lys Thr Leu Asn Gln Asn Gly Ile Arg Val Val Met Asp Leu Val Val Asn His Thr Ser Asp Thr His Ser Trp Phe Leu Asp Ala Val Glu Asn Thr Thr Asn Ser Lys Tyr Trp Ser Tyr Tyr Ile Met Thr Leu Glu Asn Arg Asp Gly Trp Asn His Trp His Trp Lys Ile Asn Ser Lys Gly Gln Lys Val Tyr Tyr Phe Gly Leu Phe Asp Ser Ser Met Pro Asp Leu Asn Phe Asp Asn Pro Gln Val Met Asn Glu Ile Lys Arg Ile Ile Asp Phe Trp Ile Thr Val Gly Val Asp Gly Phe Arg Leu Asp Ala Pro Lys His Tyr Lys Gly Trp Asp Trp Asp Asp Gly Ile Ser Gly Ser Ala Ala Ile Ala Arg Glu Ile Glu Ser Tyr Ile Arg Ser Lys Leu Gly Asn Asp Ala Ile Val Val Gly Glu Val Tyr Asp Gly Asn Pro Ser Val Leu Ser Gln Phe Ala Pro Met Pro Ala Phe Asn Phe Thr Phe Met Tyr Gly Ile Thr Gly Asn His Glu Gly Lys Asp Asn Leu Leu Gly Glu Thr Ile Ser Trp Val Asn Gly Ala Ser Tyr Tyr Leu Asn Val Lys His Phe Pro Phe Ile Asp Asn His Asp Leu Asn Arg Trp Ile Ser Ile Leu Ile Asp Gln Lys Tyr Ser Gly Asn Thr Gln Val Gly Thr Lys Gln Tyr Ile Leu Thr Asn Ala Leu Leu Leu Ser Leu Asn Gly Met Pro Val Ile Tyr Tyr Gly Asn Glu Ile Gly Leu Arg Gly Trp Lys Trp Gly Gln Asp Pro Trp Asp Leu Pro Val Arg Glu Pro Met Gln Trp Tyr Ala Ser Gln Ser Gly Ala Gly Gln Thr Trp Trp Thr Lys Pro

FIGURE 16MM

Val Tyr Gln Gln Lys Gly Ile Thr Phe Gly Asn Ala Asn Val Asp Gly Ala Met Tyr Asp Asp Pro Asn Asp Gly Val Ser Val Glu Glu Gln Met Asn Gly Tyr Thr Ile Asn Asn Phe Phe Lys Gln Phe Ile Thr Leu Arg Lys Thr Tyr Pro Ala Leu Ser Lys Gly Ser Ile Thr Ile Glu Arg Asp Trp Lys Asn Leu Tyr Val Ile Lys Arg Val Tyr Gly Asn Gln Glu Val Leu Val Leu Ile Asn Leu Asp Pro Thr Trp Pro Asn Asn Tyr Thr Leu Pro Gly Gly Tyr Arg Trp Val Trp Tyr Ala Phe Phe Asn Gly Ser Leu Phe Glu Phe Gly Asn Lys Asn Glu Ser Pro Leu Ser Gln Asp Thr Asn Trp Thr Val Asn Pro Arg Gln Val Tyr Val Phe Val Lys Asp

SEQ ID NO: 101

ttgcgattctttccaaagttaatatccccttttccgcaaaacaccagagagtggcagcgaagcgcagtatcaagagacactgaacaattacaaag
gaaagtaataatgatcaatttgaaaaaaaacaccattagcgccctggtcgcaggtatggtattaggctttgcatccaacgcaatggcggttcctag
aaccgcttttgtacacctctttgaatggaaatgggaagatgttgcacaggagtgtgaaacatttctcggacctaaaggctttgccgcagtgcaagt
ctctccgccaactaaatctcacaacacggatgcatggtggggccgttatcaacccgttagttatgcttttgaaggacgcagcggtaatcgcagcc
aatttaaaaatatggtgcaacgttgtaaagctgtaggcgtcgatatatacgtagatgcagtgattaaccacatggcagcctacgacagaaatttcc
ctgatgtaccctatagcagtaatgactttaactcctgtacaggagatattgactataataaccgttggcaaacacagcattgtgatttagtcggtctta
atgatctaaaaacaggatctgactacgtccgccaaaaaatagcggattatatgaacgacgcaatcagtatgggtgtagctggtttccgtattgatg
cagccaaacatataccagcaggtgatatagctgccattaaaggtaaattaaatggtaatccatacatcttccaagaggtaattggtgcatccggcg
aacctgttcgaccgactgaatacacctttatcggtggtgtcacggaatttcaatttgctcgaaaattgggtccagccttccgcaatagtaatattgctt
ggttaaaagacattggcagtcaaatggaattatccagtgctgatgccgtaacatttgtaacgaatcatgatgaagagcgtcataacccgaatggtc
ctatttggcacggcgttcaaggtaatggttatgcattagcaaatattttcaccttagcttacccttacggctatccaaaaatcatgtcaggatacttctt
ccacggtgactttaacgcagctccaccaagcagtggtatacacacaggaaatgcgtgtggttttgatggcggagactgggtatgcgaacacaa
atggcgcggtattgctaacatggttgccttccgcaactatacagcaagcgaatggcgtatcagtaattggtggcaaaacagtaacgaccaaattg
cttttggtcgcggtggtttaggttttgttgttattaataaacgtgctaatggtagcattaatcaaagttttgatacgggaatgcctgatggccaatactgt
aacataatagaagctaactttgatgaaagcaccggccaatgtagtgcagctacagattccaacggtcaagccgttattaccgtcagtggtgggca
agctaactttaatgtagcaggcgatcatgctgctgcaattcatgttggcgcaaaaattggtgatcaatgtagtggtgatgattgcccatgtacagga
tccgattgtaataatgatcctaaacctgattttgcagtaccagcaacatcaatttgtacatcagaaaatttacctacgctatattactggggagcaca
gcctacagatagcttagcgaatgcagcttggccaggtgtcgcaatgcaaacaaatggcgactttaagtgtcatgatttaggtgtcgaactaacca
aaattaacgccatctttagtgacaatggtgcaaataaaacagctgatctaactgttactggtgcaggttgttataaagacgggacttggagcacctt
acaaaattgtggctttgaaattaccggtgcacaaaccaatccagtcggtggcgacgaagtctggtacttccgaggtactgctaatgactggggta
agcacaattagattatgacgcaactagcggtttgtattacacaatacaaagctttaatggtgaagaagcacctgcgcgttttaaaattgataatggt
agttggactgaagcttatccaacagctgattaccaagttacagataacaattcataccgcattaactttaatagcgatagcaaagcgattacagtaa
acgcacaataa

SEQ ID NO: 102

Met Arg Phe Phe Pro Lys Leu Ile Ser Pro Phe Pro Gln Asn Thr Arg Glu Trp Gln Arg Ser Ala Val Ser Arg Asp Thr Glu Gln Leu Gln Arg Lys Val Ile Met Ile Asn Leu Lys Lys Asn Thr Ile Ser Ala Leu Val Ala Gly Met Val Leu Gly Phe Ala Ser Asn Ala Met Ala Val Pro Arg Thr Ala Phe Val His Leu Phe Glu Trp Lys Trp Glu Asp Val Ala Gln Glu Cys Glu Thr Phe Leu Gly Pro Lys Gly Phe Ala Ala Val Gln Val Ser Pro Pro Thr Lys Ser His Asn Thr Asp Ala Trp Trp Gly Arg Tyr Gln Pro Val Ser Tyr Ala Phe Glu Gly Arg Ser Gly Asn Arg Ser Gln Phe Lys Asn Met Val Gln Arg Cys Lys Ala Val Gly Val Asp Ile Tyr Val Asp Ala Val Ile Asn His Met Ala Ala Tyr Asp Arg Asn Phe Pro Asp Val Pro Tyr Ser Ser Asn Asp Phe Asn Ser Cys Thr Gly Asp Ile Asp Tyr Asn Asn Arg Trp Gln Thr Gln His Cys Asp Leu Val Gly Leu Asn Asp Leu Lys Thr Gly Ser Asp Tyr Val Arg Gln Lys Ile Ala Asp Tyr Met Asn Asp Ala Ile Ser Met Gly Val Ala Gly Phe Arg Ile Asp Ala Ala Lys His Ile Pro Ala Gly Asp Ile Ala Ala Ile Lys Gly Lys Leu Asn Gly Asn Pro Tyr Ile Phe Gln Glu Val Ile Gly Ala Ser Gly Glu Pro Val Arg Pro Thr Glu Tyr Thr Phe Ile Gly Gly Val Thr Glu Phe Gln Phe Ala Arg Lys Leu Gly Pro Ala Phe Arg Asn Ser Asn Ile Ala Trp Leu Lys Asp Ile Gly Ser Gln Met Glu Leu Ser Ser Ala Asp Ala Val Thr Phe Val Thr Asn His Asp Glu Glu Arg His Asn Pro Asn Gly Pro Ile Trp His Gly Val Gln Gly Asn Gly Tyr Ala Leu Ala Asn Ile Phe Thr Leu Ala Tyr Pro Tyr Gly Tyr Pro Lys Ile Met Ser Gly Tyr Phe Phe His Gly Asp Phe Asn Ala Ala Pro Pro Ser Ser Gly Ile His Thr

FIGURE 16NN

Gly Asn Ala Cys Gly Phe Asp Gly Gly Asp Trp Val Cys Glu His Lys Trp Arg Gly Ile Ala Asn Met Val Ala Phe Arg Asn Tyr Thr Ala Ser Glu Trp Arg Ile Ser Asn Trp Trp Gln Asn Ser Asn Asp Gln Ile Ala Phe Gly Arg Gly Gly Leu Gly Phe Val Val Ile Asn Lys Arg Ala Asn Gly Ser Ile Asn Gln Ser Phe Asp Thr Gly Met Pro Asp Gly Gln Tyr Cys Asn Ile Ile Glu Ala Asn Phe Asp Glu Ser Thr Gly Gln Cys Ser Ala Ala Thr Asp Ser Asn Gly Gln Ala Val Ile Thr Val Ser Gly Gly Gln Ala Asn Phe Asn Val Ala Gly Asp His Ala Ala Ala Ile His Val Gly Ala Lys Ile Gly Asp Gln Cys Ser Gly Asp Asp Cys Pro Cys Thr Gly Ser Asp Cys Asn Asn Asp Pro Lys Pro Asp Phe Ala Val Pro Ala Thr Ser Ile Cys Thr Ser Glu Asn Leu Pro Thr Leu Tyr Tyr Trp Gly Ala Gln Pro Thr Asp Ser Leu Ala Asn Ala Ala Trp Pro Gly Val Ala Met Gln Thr Asn Gly Asp Phe Lys Cys His Asp Leu Gly Val Glu Leu Thr Lys Ile Asn Ala Ile Phe Ser Asp Asn Gly Ala Asn Lys Thr Ala Asp Leu Thr Val Thr Gly Ala Gly Cys Tyr Lys Asp Gly Thr Trp Ser Thr Leu Gln Asn Cys Gly Phe Glu Ile Thr Gly Ala Gln Thr Asn Pro Val Gly Gly Asp Glu Val Trp Tyr Phe Arg Gly Thr Ala Asn Asp Trp Gly Lys Ala Gln Leu Asp Tyr Asp Ala Thr Ser Gly Leu Tyr Tyr Thr Ile Gln Ser Phe Asn Gly Glu Glu Ala Pro Ala Arg Phe Lys Ile Asp Asn Gly Ser Trp Thr Glu Ala Tyr Pro Thr Ala Asp Tyr Gln Val Thr Asp Asn Asn Ser Tyr Arg Ile Asn Phe Asn Ser Asp Ser Lys Ala Ile Thr Val Asn Ala Gln

SEQ ID NO: 103

gtgctaacgtttcaccgcatcattcgaaaaggatggatgttcctgctcgcgtttttgctcactgcctcgctgttctgcccaacaggacagcccgcca
aggctgccgcaccgtttaacggcaccatgatgcagtattttgaatggtacttgccggatgatggcacgttatggaccaaagtggccaatgaagc
caacaacttatccagccttggcatcaccgctctttggctgccgcccgcttacaaaggaacaagccgcagcgacgtagggtacggagtatacga
cttgtatgacctcggcgaattcaatcaaaaagggaccgtccgcacaaaatacggaacaaaagctcaatatcttcaagccattcaagccgcccac
gccgctggaatgcaagtgtacgccgatgtcgtgttcgaccataaggcggcgccgacggcacggaatgggtggacgccgtcgaagtcaatc
cgtccgaccgcaaccaagaaatctcgggcacctatcaaatccaagcatggacgaaatttgattttcccgggcggggcaacacctactccagctt
taagtggcgctggtaccattttgacggcgttgattgggacgaaagccgaaaattgagccgcatttacaaattccgcggcatcggcaaagcgtgg
gattgggaagtagacacggaaaacggaaactatgactacttaatgtatgccgaccttgatatggatcatcccgaagtcgtgaccgagctgaaaa
actgggggaatggtatgtcaacacaacgaacattgatgggttccggcttgatgccgtcaagcatattaagttcagttttttcctgattggttgtcgt
atgtgcgttctcagactggcaagccgctatttaccgtcggggaatattggagctatgacatcaacaagttgcacaattacattacgaaaacaaacg
gaacgatgtctttgtttgatgccccgttacacaacaaattttataccgcttccaaatcaggggcgcatttgatatgcgcacgttaatgaccaatact
ctcatgaaagatcaaccgacattggccgtcaccttcgttgataatcatgacaccgaacccggccaagcgctgcagtcatgggtcgacccatggt
tcaaaccgttggcttacgcctttattctaactcggcaggaaggatacccgtgcgtcttttatggtgactattatggcattccacaatataacattccttc
gctgaaaagcaaaatcgatccgctcctcatcgcgcgcagggattatgcttacggaacgcaacatgattatcttgatcactccgacatcatcgggt
ggacaagggaaggggtcactgaaaaaccaggatccgggctggccgcactgatcacgatgggccgggaggaagcaaatggatgtacgttg
gcaaacaacacgctggaaaagtgttctatgaccttaccggcaaccggagtgacaccgtcaccatcaacagtgatggatgggggaattcaaa
gtcaatggcggttcggtttcggtttgggttcctagaaaaacgaccgtttctaccatcgctcggccgatcacaacccgaccgtggactggtgaattc
gtccgttggaccgaaccacggttggtggcatggccttga

SEQ ID NO: 104

Val Leu Thr Phe His Arg Ile Ile Arg Lys Gly Trp Met Phe Leu Leu Ala Phe Leu Leu Thr Ala Ser Leu Phe Cys Pro Thr Gly Gln Pro Ala Lys Ala Ala Ala Pro Phe Asn Gly Thr Met Met Gln Tyr Phe Glu Trp Tyr Leu Pro Asp Asp Gly Thr Leu Trp Thr Lys Val Ala Asn Glu Ala Asn Asn Leu Ser Ser Leu Gly Ile Thr Ala Leu Trp Leu Pro Pro Ala Tyr Lys Gly Thr Ser Arg Ser Asp Val Gly Tyr Gly Val Tyr Asp Leu Tyr Asp Leu Gly Glu Phe Asn Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly Thr Lys Ala Gln Tyr Leu Gln Ala Ile Gln Ala Ala His Ala Ala Gly Met Gln Val Tyr Ala Asp Val Val Phe Asp His Lys Gly Gly Ala Asp Gly Thr Glu Trp Val Asp Ala Val Glu Val Asn Pro Ser Asp Arg Asn Gln Glu Ile Ser Gly Thr Tyr Gln Ile Gln Ala Trp Thr Lys Phe Asp Phe Pro Gly Arg Gly Asn Thr Tyr Ser Ser Phe Lys Trp Arg Trp Tyr His Phe Asp Gly Val Asp Trp Asp Glu Ser Arg Lys Leu Ser Arg Ile Tyr Lys Phe Arg Gly Ile Gly Lys Ala Trp Asp Trp Glu Val Asp Thr Glu Asn Gly Asn Tyr Asp Tyr Leu Met Tyr Ala Asp Leu Asp Met Asp His Pro Glu Val Val Thr Glu Leu Lys Asn Trp Gly Glu Trp Tyr Val Asn Thr Thr Asn Ile Asp Gly Phe Arg Leu Asp Ala Val Lys His Ile Lys Phe Ser Phe Phe Pro Asp Trp Leu Ser Tyr Val Arg Ser Gln Thr Gly Lys Pro Leu Phe Thr Val Gly

FIGURE 16OO

Glu Tyr Trp Ser Tyr Asp Ile Asn Lys Leu His Asn Tyr Ile Thr Lys Thr Asn Gly Thr Met Ser Leu Phe Asp Ala Pro Leu His Asn Lys Phe Tyr Thr Ala Ser Lys Ser Gly Gly Ala Phe Asp Met Arg Thr Leu Met Thr Asn Thr Leu Met Lys Asp Gln Pro Thr Leu Ala Val Thr Phe Val Asp Asn His Asp Thr Glu Pro Gly Gln Ala Leu Gln Ser Trp Val Asp Pro Trp Phe Lys Pro Leu Ala Tyr Ala Phe Ile Leu Thr Arg Gln Glu Gly Tyr Pro Cys Val Phe Tyr Gly Asp Tyr Tyr Gly Ile Pro Gln Tyr Asn Ile Pro Ser Leu Lys Ser Lys Ile Asp Pro Leu Leu Ile Ala Arg Arg Asp Tyr Ala Tyr Gly Thr Gln His Asp Tyr Leu Asp His Ser Asp Ile Ile Gly Trp Thr Arg Glu Gly Val Thr Glu Lys Pro Gly Ser Gly Leu Ala Ala Leu Ile Thr Asp Gly Pro Gly Gly Ser Lys Trp Met Tyr Val Gly Lys Gln His Ala Gly Lys Val Phe Tyr Asp Leu Thr Gly Asn Arg Ser Asp Thr Val Thr Ile Asn Ser Asp Gly Trp Gly Glu Phe Lys Val Asn Gly Gly Ser Val Ser Val Trp Val Pro Arg Lys Thr Thr Val Ser Thr Ile Ala Arg Pro Ile Thr Thr Arg Pro Trp Thr Gly Glu Phe Val Arg Trp Thr Glu Pro Arg Leu Val Ala Trp Pro

SEQ ID NO: 105

atgtccctattcaaaaaaatctttccgtggattgtatctctacttcttttgttttcgtttattgctccttttttccattcaaacagaaaaagtccgcgctggaa
gtgttccagtgaatggaacgatgatgcaatatttcgaatggtaccttccagacgatggaacactatggacgaaagtagcaaataacgcccaatct
ttagcgaatcttggcattactgccctttggcttccccctgcctataaaggaacaagcagcagtgacgttggatatggcgtttatgatttatatgacct
aggagagtttaatcaaaaaggaactgtccgaacaaaatacggaacaaaaacacaatatatccaagcaatccaagcggcgcatacagcaggaa
tgcaagtatatgcagatgtcgtctttaaccataaagccggtgcagatgggacagaactagtggatgcagtagaagtaaacccttctgaccgcaat
caagaaatatcaggaacatatcaaatccaagcgtggacaaaatttgattttcctggtcgtggaaacacctattctagttttaaatggcgttggtatca
tttcgatggaacggactgggatgagagtagaaaactaaatcgtatttacaaattccgcggcacgggaaaagcatgggattgggaagtagataca
gaaaatgggaattatgactatctcatgtatgcagatttggatatggatcatccagaggttgtatctgaactaaaaaattggggaaagtggtatgtaa
ccacaaccaatatcgacggattccgtctggatgcagtgaagcatattaaatatagcttttcccagactggctatcgtatgtacgaacccaaacac
aaaagcctcttttgccgttggcgaattttggagctatgacattaacaagctacacaactatattacaaagacgaacggctctatgtccctattcgat
gccccgctgcataacaatttttatatagcatcgaaatcaggtggctattttgatatgcgcacattactcaacaacacattgatgaaagatcaaccaa
cactatcggtcacattagtagacaatcacgatactgagccagggcaatctttgcagtcgtgggtcgagccgtggtttaaaccgttagcttacgcat
ttatcttgacccgccaagaaggttatccgtgcatcttttatggagattactatggtattccaaaatacaacattcctgcgctgaaaagcaaacttgatc
cgctgttaattgctcgaagagattatgcctacggaacacagcacgactatattgacaatgcagatattatcggctggacgcgggaaggagtagct
gaaaaagcaaattcgggacttgctgcactcattaccgacggacctggcggaagcaaatggatgtatgttggcaaacaacacgctggcaaaac
gttttatgatctaaccggcaatcgaagtgatacagtgacaatcaacgctgatggatggggagaatttaaagtcaatggagggtctgtatccatatg
ggttccaaaaacatcaaccacttcccaaatcacatttactgtaaataatgccacaaccgtttggggacaaaatgtatacgttgtcgggaatatttcg
cagctgggcaac

SEQ ID NO: 106

Met Ser Leu Phe Lys Lys Ile Phe Pro Trp Ile Val Ser Leu Leu Leu Leu Phe Ser Phe Ile Ala Pro Phe Ser Ile Gln Thr Glu Lys Val Arg Ala Gly Ser Val Pro Val Asn Gly Thr Met Met Gln Tyr Phe Glu Trp Tyr Leu Pro Asp Asp Gly Thr Leu Trp Thr Lys Val Ala Asn Asp Ala Glu Ser Leu Ala Asn Leu Gly Ile Thr Ala Leu Trp Leu Pro Pro Ala Tyr Lys Gly Thr Ser Ser Ser Asp Val Gly Tyr Gly Val Tyr Asp Leu Tyr Asp Leu Gly Glu Phe Asn Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly Thr Lys Thr Gln Tyr Ile Gln Ala Ile Gln Ala Ala His Thr Ala Gly Met Gln Val Tyr Ala Asp Val Val Phe Asn His Lys Ala Gly Ala Asp Gly Thr Glu Leu Val Asp Ala Val Glu Val Asn Pro Ser Asp Arg Asn Gln Glu Ile Ser Gly Thr Tyr Gln Ile Gln Ala Trp Thr Lys Phe Asp Phe Pro Gly Arg Gly Asn Thr Tyr Ser Ser Phe Lys Trp Arg Trp Tyr His Phe Asp Gly Thr Asp Trp Asp Glu Ser Arg Lys Leu Asn Arg Ile Tyr Lys Phe Arg Gly Thr Gly Lys Ala Trp Asp Trp Glu Val Asp Thr Glu Asn Gly Asn Tyr Asp Tyr Leu Met Tyr Ala Asp Leu Asp Met Asp His Pro Glu Val Val Ser Glu Leu Lys Asn Trp Gly Lys Trp Tyr Val Thr Thr Thr Asn Ile Asp Gly Phe Arg Leu Asp Ala Val Lys His Ile Lys Tyr Ser Phe Phe Pro Asp Trp Leu Ser Tyr Val Arg Thr Gln Thr Gln Lys Pro Leu Phe Ala Val Gly Glu Phe Trp Ser Tyr Asp Ile Asn Lys Leu His Asn Tyr Ile Thr Lys Thr Asn Gly Ser Met Ser Leu Phe Asp Ala Pro Leu His Asn Asn Phe Tyr Ile Ala Ser Lys Ser Gly Gly Tyr Phe Asp Met Arg Thr Leu Leu Asn Asn Thr Leu Met Lys Asp Gln Pro Thr Leu Ser Val Thr Leu Val Asp Asn His Asp Thr

FIGURE 16PP

Glu Pro Gly Gln Ser Leu Gln Ser Trp Val Glu Pro Trp Phe Lys Pro Leu Ala Tyr Ala Phe Ile Leu Thr Arg Gln Glu Gly Tyr Pro Cys Ile Phe Tyr Gly Asp Tyr Tyr Gly Ile Pro Lys Tyr Asn Ile Pro Ala Leu Lys Ser Lys Leu Asp Pro Leu Leu Ile Ala Arg Arg Asp Tyr Ala Tyr Gly Thr Gln His Asp Tyr Ile Asp Asn Ala Asp Ile Ile Gly Trp Thr Arg Glu Gly Val Ala Glu Lys Ala Asn Ser Gly Leu Ala Ala Leu Ile Thr Asp Gly Pro Gly Gly Ser Lys Trp Met Tyr Val Gly Lys Gln His Ala Gly Lys Thr Phe Tyr Asp Leu Thr Gly Asn Arg Ser Asp Thr Val Thr Ile Asn Ala Asp Gly Trp Gly Glu Phe Lys Val Asn Gly Gly Ser Val Ser Ile Trp Val Pro Lys Thr Ser Thr Thr Ser Gln Ile Thr Phe Thr Val Asn Asn Ala Thr Thr Val Trp Gly Gln Asn Val Tyr Val Val Gly Asn Ile Ser Gln Leu Gly Asn

SEQ ID NO: 107

atggacagcctcgacgcgccggagcagaagccctgggtgaaggatggcaggctctccgcgtacctggatacagggacagggaccgtggtc
gctcccgaggcacctgcgcccccgccgcccccggccgaggaagtccggcccgtggacaagtggaaaaacgatatcatctatttcgtcctcac
cgaccgtttccaggatggcgacaagaccaacaacatggacgtggtcccgacggacatgaaaaaatatcatggcggcgacatccaggggctc
atcgacaagctcgactatatcaaggagaccggttcgacggccatctggctcacgccccctatgaaggggcagacccacttcttcgagaccgac
aattaccatggttactggcccattgacttctatgacacggaccccccatgtgggcaccatgcagaaatttgaggagcttatcgagaaagcccatga
gaaagggctgaagatcgtgctcgatattcccctgaaccacacggcctgggagcatccccttctacaaggacgacagcaagaaggactggttcc
accatataggagatgtgaaggactgggaagatccctactgggctgaaaacggctccatattcggtcttcctgacctggcgcaggaaaaccctg
ccgtggaaaagtacctcatcgacgtggccaagttctgggtagacaagggtattgacggcttcaggcttgacgccgtgaagaacgtgccccctca
acttctgggcgaagtttgaccgggcgattcacgattatgcgggcaaggacttcctcctcgtcggggaatactttgacggaaacccggcgaaagt
cgcgaactaccagagagaggacatgagctcactcttcgattacccgctctactggaccctgaaggacaccttcgccaaggacgggagcatgc
gcaacctggcggcgaagcttgatgagtgcgacaggaattatcccgacccgggcctcatgtcggttttccttgataaccacgacacgccgaggtt
cctcaccgaggccaacggcaacaaggataagctcaaactggccctcgccttcgcgatgaccatcaaccgcatgcctaccatttattatggcacc
gaggttgccatggaaggcaactgcgatatcatgggcgccgtagataaccggagggacatgcagtgggacaaggatcctgacatgttcaaata
cttcaagactctcaccactgcccgcaatgagcatgaatccctcagggaaggaaagaagctcgagatgtggcaggatgacaaagtctacgcgta
cgggaggcagaccccgaaggacgagtctatcgtggtgcttaacaacggctatgatacgcaggaacgggacataccgctccgccccgagag
cggcatcaagaacggcacggtgctgaaggatgtcatcaccggcgaaaccgtgacggtacagaacggaaaaatccatgcgaaatgcggcgg
caaacaggcgcggatctacgtgcccgcgtag

SEQ ID NO: 108

Met Asp Ser Leu Asp Ala Pro Glu Gln Lys Pro Trp Val Lys Asp Gly Arg Leu Ser Ala Tyr Leu Asp Thr Gly Thr Gly Thr Val Val Ala Pro Glu Ala Pro Ala Pro Pro Pro Pro Pro Ala Glu Glu Val Arg Pro Val Asp Lys Trp Lys Asn Asp Ile Ile Tyr Phe Val Leu Thr Asp Arg Phe Gln Asp Gly Asp Lys Thr Asn Asn Met Asp Val Val Pro Thr Asp Met Lys Lys Tyr His Gly Gly Asp Ile Gln Gly Leu Ile Asp Lys Leu Asp Tyr Ile Lys Glu Thr Gly Ser Thr Ala Ile Trp Leu Thr Pro Pro Met Lys Gly Gln Thr His Phe Phe Glu Thr Asp Asn Tyr His Gly Tyr Trp Pro Ile Asp Phe Tyr Asp Thr Asp Pro His Val Gly Thr Met Gln Lys Phe Glu Glu Leu Ile Glu Lys Ala His Glu Lys Gly Leu Lys Ile Val Leu Asp Ile Pro Leu Asn His Thr Ala Trp Glu His Pro Phe Tyr Lys Asp Asp Ser Lys Lys Asp Trp Phe His His Ile Gly Asp Val Lys Asp Trp Glu Asp Pro Tyr Trp Ala Glu Asn Gly Ser Ile Phe Gly Leu Pro Asp Leu Ala Gln Glu Asn Pro Ala Val Glu Lys Tyr Leu Ile Asp Val Ala Lys Phe Trp Val Asp Lys Gly Ile Asp Gly Phe Arg Leu Asp Ala Val Lys Asn Val Pro Leu Asn Phe Trp Ala Lys Phe Asp Arg Ala Ile His Asp Tyr Ala Gly Lys Asp Phe Leu Leu Val Gly Glu Tyr Phe Asp Gly Asn Pro Ala Lys Val Ala Asn Tyr Gln Arg Glu Asp Met Ser Ser Leu Phe Asp Tyr Pro Leu Tyr Trp Thr Leu Lys Asp Thr Phe Ala Lys Asp Gly Ser Met Arg Asn Leu Ala Ala Lys Leu Asp Glu Cys Asp Arg Asn Tyr Pro Asp Pro Gly Leu Met Ser Val Phe Leu Asp Asn His Asp Thr Pro Arg Phe Leu Thr Glu Ala Asn Gly Asn Lys Asp Lys Leu Lys Leu Ala Leu Ala Phe Ala Met Thr Ile Asn Arg Met Pro Thr Ile Tyr Tyr Gly Thr Glu Val Ala Met Glu Gly Asn Cys Asp Ile Met Gly Ala Val Asp Asn Arg Arg Asp Met Gln Trp Asp Lys Asp Pro Asp Met Phe Lys Tyr Phe Lys Thr Leu Thr Thr Ala Arg Asn Glu His Glu Ser Leu Arg Glu Gly Lys Lys Leu Glu Met Trp Gln Asp Asp Lys Val Tyr Ala Tyr Gly Arg Gln Thr Pro Lys Asp Glu Ser Ile Val Val Leu Asn Asn Gly Tyr Asp Thr Gln Glu Arg Asp Ile

FIGURE 16QQ

Pro Leu Arg Pro Glu Ser Gly Ile Lys Asn Gly Thr Val Leu Lys Asp Val Ile Thr Gly Glu Thr Val Thr Val Gln Asn Gly Lys Ile His Ala Lys Cys Gly Gly Lys Gln Ala Arg Ile Tyr Val Pro Ala

SEQ ID NO: 109

atggcaagaaaaacgctggccatatttttcgtacttctagtgcttcttagtctctcggcagttccggcaaaggcagaaactctagagaatggtgga
gttataatgcaggctttctattgggatgttcctggaggaggaatctggtgggacacaatagctcaaaagataccccgaatgggcaagtgcaggaat
ctcagcgatatggattccaccagcgagtaagggcatgagcggtggttattccatgggctacgatccctacgatttctttgacctcggcgagtacta
tcagaaggggacagttgagacgcgcttcggctcaaaggaagaactggtgaacatgataaacaccgcacactcctacggcataaaggtgatag
cggacatagtcataaaccaccgcgccggtggagaccttgagtggaaccccttcgtgaacgactatacctggacagacttctcaaaagtcgcctc
cggtaaatatacggccaactaccttgacttccacccaaacgagcttcactgttgtgatgaaggtacctttggaggataccctgatatatgtcacga
caaaagctgggaccagtactggctctgggcgagcagcgaaagctacgctgcctacctcaggagcatagggttgacgcctggcgtttcgact
acgtcaagggctacggagcatgggttgttaacgactggctcagctggtggggaggctgggccgttggagagtactgggacacgaacgttgat
gcactcctcaactgggcatacagcagcggcgccaaggtctttgacttcccgctctactacaagatggacgaagccttcgacaacaccaacatcc
cggcattagtggatgcactcagatacggccagacagtggtcagccgcgatcccttcaaggcggtaactttcgttgccaaccacgatacagatat
aatctggaacaagtatccggcttatgcattcatccttacctatgagggacagcctgttatattctaccgcgactacgaggagtggctcaacaagga
taagcttaacaacctcatctggatacacgatcaccttgctggagggagtactgacattgtttactacgacagcgacgagcttatctttgtgagaaac
ggctatggcaccaaaccaggactgataacctatatcaacctcggctcaagcaaagttggaaggtgggtctacgttccaaagttcgccggttcat
gcatccacgagtacaccggcaacctcggcggttggatagacaagtacgtctcctccagcggctgggtctatcttgaggccccagcccacgac
ccggcgaacggctactacggctactctgtctggagctactgcggtgtgggttga

SEQ ID NO: 110

Met Ala Arg Lys Thr Leu Ala Ile Phe Phe Val Leu Leu Val Leu Leu Ser Leu Ser Ala Val Pro Ala Lys Ala Glu Thr Leu Glu Asn Gly Gly Val Ile Met Gln Ala Phe Tyr Trp Asp Val Pro Gly Gly Gly Ile Trp Trp Asp Thr Ile Ala Gln Lys Ile Pro Glu Trp Ala Ser Ala Gly Ile Ser Ala Ile Trp Ile Pro Pro Ala Ser Lys Gly Met Ser Gly Gly Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Phe Phe Asp Leu Gly Glu Tyr Tyr Gln Lys Gly Thr Val Glu Thr Arg Phe Gly Ser Lys Glu Glu Leu Val Asn Met Ile Asn Thr Ala His Ser Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Asn Asp Tyr Thr Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr Leu Asp Phe His Pro Asn Glu Leu His Cys Cys Asp Glu Gly Thr Phe Gly Gly Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp Leu Trp Ala Ser Ser Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly Val Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp Val Val Asn Asp Trp Leu Ser Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn Val Asp Ala Leu Leu Asn Trp Ala Tyr Ser Ser Gly Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Glu Ala Phe Asp Asn Thr Asn Ile Pro Ala Leu Val Asp Ala Leu Arg Tyr Gly Gln Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile Leu Thr Tyr Glu Gly Gln Pro Val Ile Phe Tyr Arg Asp Tyr Glu Glu Trp Leu Asn Lys Asp Lys Leu Asn Asn Leu Ile Trp Ile His Asp His Leu Ala Gly Gly Ser Thr Asp Ile Val Tyr Tyr Asp Ser Asp Glu Leu Ile Phe Val Arg Asn Gly Tyr Gly Thr Lys Pro Gly Leu Ile Thr Tyr Ile Asn Leu Gly Ser Ser Lys Val Gly Arg Trp Val Tyr Val Pro Lys Phe Ala Gly Ser Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp Ile Asp Lys Tyr Val Ser Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro Ala His Asp Pro Ala Asn Gly Tyr Tyr Gly Tyr Ser Val Trp Ser Tyr Cys Gly Val Gly

SEQ ID NO: 111

atgcccgcgttcaaatctaaggtgatgcacatgaagttgaagtaccttgctttagttttgttggctgtggcttcgataggcctcctctcgactccagt
gggtgctgccaagtactccgaactcgaagagggcggtgttataatgcaggcttctactgggacgtccctaccggtgggatctggtgggacac
cataagacagaaaatcccggagtggtacgacgctggaatctcggcgatatggattcctccagctagcaaaggtatgggtggtgcatactccatg
ggttatgacccctacgatttctttgacctcggcgagtactatcagaagggaacagttgagacgcgcttcggctcaaaggaggaactggtgaaca
tgataaacaccgcacactcctatggcataaaggtgatagcggacatagtcataaaccaccgcgccggcggcgacctggagtggaacccctttg
taaacaactatacttggacagacttctccaaggtcgcctccggtaaatacacggccaactaccttgacttccacccaaacgaggtcaagtgctgc
gatgagggtacatttggtgactttccggacatcgcccacgagaagagctgggatcagtactggctctgggcaagcaatgagagctacgccgcc

2013201807 18 Mar 2013

FIGURE 16RR

tatctccggagcatagggatcgatgcatggcgtttcgactacgtcaaaggttacggagcgtgggttgttaacgactggctcagctggtggggag
gttgggccgttggagagtactgggacaccaacgttgatgcactccttaactgggcatacaacagcggtgccaaggtctttgacttccccgctctac
tacaagatggacgaagcctttgacaacaccaacatccccgctttggtttacgccctccagaacggaggaacagtcgtttcccgcgatcccttcaa
ggcagtaactttcgttgccaaccacgataccgatataatctggaacaagtatccggcttatgcgttcatccttacctatgagggacagcctgttatat
tctaccgcgactacgaggagtggctcaacaaggataagcttaacaaccttatctggatacacgagcaccttgccggaggaagtaccaagatcct
ctactacgataacgatgagctaatattcatgagggagggctacgggagcaagccgggcctcataacctacataaacctcggaaacgactggg
ccgagcgctgggtgaacgtcggctcaaagtttgccggctacacaatccatgaatacacaggcaatctcggtggctgggttgacaggtgggttc
agtacgacggatgggttaaactgacggcacctcctcacgatccagccaacggatattacggctactcagtctggagctacgcaggcgtcggat
ga

SEQ ID NO: 112
Met Pro Ala Phe Lys Ser Lys Val Met His Met Lys Leu Lys Tyr Leu Ala Leu Val Leu Leu Ala Val
Ala Ser Ile Gly Leu Leu Ser Thr Pro Val Gly Ala Ala Lys Tyr Ser Glu Leu Glu Glu Gly Gly Val Ile
Met Gln Ala Phe Tyr Trp Asp Val Pro Thr Gly Gly Ile Trp Trp Asp Thr Ile Arg Gln Lys Ile Pro
Glu Trp Tyr Asp Ala Gly Ile Ser Ala Ile Trp Ile Pro Pro Ala Ser Lys Gly Met Gly Gly Ala Tyr Ser
Met Gly Tyr Asp Pro Tyr Asp Phe Phe Asp Leu Gly Glu Tyr Tyr Gln Lys Gly Thr Val Glu Thr
Arg Phe Gly Ser Lys Glu Glu Leu Val Asn Met Ile Asn Thr Ala His Ser Tyr Gly Ile Lys Val Ile
Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Asn Asn Tyr Thr
Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr Leu Asp Phe His Pro Asn Glu
Val Lys Cys Cys Asp Glu Gly Thr Phe Gly Asp Phe Pro Asp Ile Ala His Glu Lys Ser Trp Asp Gln
Tyr Trp Leu Trp Ala Ser Asn Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly Ile Asp Ala Trp Arg
Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp Val Val Asn Asp Trp Leu Ser Trp Trp Gly Gly Trp Ala
Val Gly Glu Tyr Trp Asp Thr Asn Val Asp Ala Leu Leu Asn Trp Ala Tyr Asn Ser Gly Ala Lys Val
Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Glu Ala Phe Asp Asn Thr Asn Ile Pro Ala Leu Val Tyr
Ala Leu Gln Asn Gly Gly Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn His
Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile Leu Thr Tyr Glu Gly Gln Pro Val Ile
Phe Tyr Arg Asp Tyr Glu Glu Trp Leu Asn Lys Asp Lys Leu Asn Asn Leu Ile Trp Ile His Glu His
Leu Ala Gly Gly Ser Thr Lys Ile Leu Tyr Tyr Asp Asn Asp Glu Leu Ile Phe Met Arg Glu Gly Tyr
Gly Ser Lys Pro Gly Leu Ile Thr Tyr Ile Asn Leu Gly Asn Asp Trp Ala Glu Arg Trp Val Asn Val
Gly Ser Lys Phe Ala Gly Tyr Thr Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp Val Asp Arg Trp
Val Gln Tyr Asp Gly Trp Val Lys Leu Thr Ala Pro Pro His Asp Pro Ala Asn Gly Tyr Tyr Gly Tyr
Ser Val Trp Ser Tyr Ala Gly Val Gly

SEQ ID NO: 113
atgaaacaacaaaaacggctttacgcccgattgctgacgctgttatttgcgctcatcttcttgctgcctcattctgcagcagcggcggcaaatctta
atgggacgctgatgcagtatttgaatggtacatgcccaatgacggccaacattggaagcgcttgcaaaacgactcggcatatttggctgaacac
ggtattactgccgtctggattcccccggcatataagggaacgagccaagcggatgtgggctacggtgcttacgacctttatgatttaggggagttt
catcaaaaagggacggttcggacaaagtacggcacaaaaggagagctgcaatctgcgatcaaaagtcttcattcccgcgacattaacgtttacg
gggatgtggtcatcaaccacaaaggcggcgctgatgcgaccgaagatgtaaccgcggttgaagtcgatcccgctgaccgcaaccgcgtaatt
tcaggagaacaccgaattaaagcctggacacatttcatttccggggcgcggcagcacatacagcgatttaaatggcattggtaccattttgac
ggaaccgattgggacgagtcccgaaagctgaaccgcatctataagtttcaaggaaaggcttgggattgggaagtttccaatgaaaacggcaac
tatgattatttgatgtatgccgacatcgattatgaccatcctgatgtcgcagcagaaattaagagatggggcacttggtatgccaatgaactgcaatt
ggacggtttccgtcttgatgctgtcaaacacattaaattttcttttttgcgggattgggttaatcatgtcagggaaaaaacggggaaggaaatgttta
cggtagctgaatattggcagaatgacttgggcgcgctggaaaactatttgaacaaaacaaatttaatcattcagtgtttgacgtgccgcttcattat
cagttccatgctgcatcgacacagggaggcggctatgatatgaggaaattgctgaacggtacggtcgtttccaagcatccgttgaaagcggtta
catttgtcgataaccatgatacacagccggggcaatcgcttgagtcgactgtccaaacatggtttaagccgcttgcttacgctttcattctcacaag
ggaatctggataccctcaggttttctacggggatatgtacgggacgaaaggagactcccagcgcgaaattcctgccttgaaacacaaaattgaa
ccgatcttaaaagcgagaaaacagtatgcgtacggagcacagcatgattatttcgaccaccatgacattgtcggctggacaagggaaggcgac
agctcggttgcaaattcaggtttggcggcattaataacagacggacccggtggggcaaagcgaatgtatgtcggccggcaaaacgccggtga

FIGURE 16SS

gacatggcatgacattaccggaaaccgttcggagccggttgtcatcaattcggaaggctggggagagtttcacgtaaacggcgggtcggtttca
atttatgttcaaagatag

SEQ ID NO: 114
Met Lys Gln Gln Lys Arg Leu Tyr Ala Arg Leu Leu Thr Leu Leu Phe Ala Leu Ile Phe Leu Leu Pro
His Ser Ala Ala Ala Ala Ala Asn Leu Asn Gly Thr Leu Met Gln Tyr Phe Glu Trp Tyr Met Pro Asn
Asp Gly Gln His Trp Lys Arg Leu Gln Asn Asp Ser Ala Tyr Leu Ala Glu His Gly Ile Thr Ala Val
Trp Ile Pro Pro Ala Tyr Lys Gly Thr Ser Gln Ala Asp Val Gly Tyr Gly Ala Tyr Asp Leu Tyr Asp
Leu Gly Glu Phe His Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly Thr Lys Gly Glu Leu Gln Ser Ala
Ile Lys Ser Leu His Ser Arg Asp Ile Asn Val Tyr Gly Asp Val Val Ile Asn His Lys Gly Gly Ala
Asp Ala Thr Glu Asp Val Thr Ala Val Glu Val Asp Pro Ala Asp Arg Asn Arg Val Ile Ser Gly Glu
His Arg Ile Lys Ala Trp Thr His Phe His Phe Pro Gly Arg Gly Ser Thr Tyr Ser Asp Phe Lys Trp
His Trp Tyr His Phe Asp Gly Thr Asp Trp Asp Glu Ser Arg Lys Leu Asn Arg Ile Tyr Lys Phe Gln
Gly Lys Ala Trp Asp Trp Glu Val Ser Asn Glu Asn Gly Asn Tyr Asp Tyr Leu Met Tyr Ala Asp Ile
Asp Tyr Asp His Pro Asp Val Ala Ala Glu Ile Lys Arg Trp Gly Thr Trp Tyr Ala Asn Glu Leu Gln
Leu Asp Gly Phe Arg Leu Asp Ala Val Lys His Ile Lys Phe Ser Phe Leu Arg Asp Trp Val Asn His
Val Arg Glu Lys Thr Gly Lys Glu Met Phe Thr Val Ala Glu Tyr Trp Gln Asn Asp Leu Gly Ala
Leu Glu Asn Tyr Leu Asn Lys Thr Asn Phe Asn His Ser Val Phe Asp Val Pro Leu His Tyr Gln Phe
His Ala Ala Ser Thr Gln Gly Gly Gly Tyr Asp Met Arg Lys Leu Leu Asn Gly Thr Val Val Ser Lys
His Pro Leu Lys Ala Val Thr Phe Val Asp Asn His Asp Thr Gln Pro Gly Gln Ser Leu Glu Ser Thr
Val Gln Thr Trp Phe Lys Pro Leu Ala Tyr Ala Phe Ile Leu Thr Arg Glu Ser Gly Tyr Pro Gln Val
Phe Tyr Gly Asp Met Tyr Gly Thr Lys Gly Asp Ser Gln Arg Glu Ile Pro Ala Leu Lys His Lys Ile
Glu Pro Ile Leu Lys Ala Arg Lys Gln Tyr Ala Tyr Gly Ala Gln His Asp Tyr Phe Asp His His Asp
Ile Val Gly Trp Thr Arg Glu Gly Asp Ser Ser Val Ala Asn Ser Gly Leu Ala Ala Leu Ile Thr Asp
Gly Pro Gly Gly Ala Lys Arg Met Tyr Val Gly Arg Gln Asn Ala Gly Glu Thr Trp His Asp Ile Thr
Gly Asn Arg Ser Glu Pro Val Val Ile Asn Ser Glu Gly Trp Gly Glu Phe His Val Asn Gly Gly Ser
Val Ser Ile Tyr Val Gln Arg

SEQ ID NO: 115
atggcgaagtactccgagctggagcagggcggagtcataatgcaggccttctactgggacgttccggagggaggaatctggtgggacacaat
acggcagaagatccctgaatggtacgatgcaggcatatccgccatctggataccccggcgagcaagggcatgggcggggcctactcgatg
ggctacgacccctacgattacttcgatctgggcgagttttaccagaagggaaccgttgagacccgcttcggctccaaggaagagctcgtcaaca
tgatctccacggcccaccagtacggcatcaaggttatagcggacatagtgataaaccaccgcgcaggtggagacctcgaatggaacccatac
gtcggcgactatacctggacggacttttctaaggtcgcctccgggaaatacaaggcccactacatggacttccatccaaacaactacagcacct
cagacgagggaaccttcggtggcttcccagacattgatcacctcgtgcccttcaaccagtactggctgtgggcgagcaacgagagctacgccg
cctacctcaggagcatagggatcgatgcgtggcgcttttgactacgttaagggctacggcgcgtgggtcgtcaaggactggctgagtcagtggg
gcggctggcgtcgtggaagtactggacaacaacgtcgatggcctcctaactgggcctacaggaacggtgcgccaaggttctcgacttccc
gctctactacaaggatggacgaggccttgacaacaaggacattccgcccctcgtttacgccatccagaacggtgtaaaccgtcgtcagcagggat
cccttcaaggccgttaccttcgtggctaaccacgatacgaacataatctggaacaagtaccctgcctatgccttcatcctgacctacgaaggtcag
cccgtcatcttctaccgcgactacgaggagtggctcaacaaggacaaactcaacaacctcatatggattcacgagcacctggcaggggaaag
caccaagatcctctactacgacgacgatgagctcatcttcatgagggaaggctacggcgacaggcccgggcttataacctacatcaacctcggt
agcgactgggcggagagatgggtgaacgttggctcaaagttcgcgggctatacaatccacgaatacaccggaaacctcggcggctgggtcg
acaggtacgtccagtacgacggctgggtcaagcttaccgctccgccacacgatccggcaaacggctattacggctactcggtctggagctacg
ccggagttggaagatctcatcaccatcaccatcactaa

SEQ ID NO: 116
Met Ala Lys Tyr Ser Glu Leu Glu Gln Gly Gly Val Ile Met Gln Ala Phe Tyr Trp Asp Val Pro Glu
Gly Gly Ile Trp Trp Asp Thr Ile Arg Gln Lys Ile Pro Glu Trp Tyr Asp Ala Gly Ile Ser Ala Ile Trp
Ile Pro Pro Ala Ser Lys Gly Met Gly Gly Ala Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Tyr Phe Asp
Leu Gly Glu Phe Tyr Gln Lys Gly Thr Val Glu Thr Arg Phe Gly Ser Lys Glu Glu Leu Val Asn Met

FIGURE 16TT

Ile Ser Thr Ala His Gln Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Tyr Val Gly Asp Tyr Thr Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Lys Ala His Tyr Met Asp Phe His Pro Asn Asn Tyr Ser Thr Ser Asp Glu Gly Thr Phe Gly Gly Phe Pro Asp Ile Asp His Leu Val Pro Phe Asn Gln Tyr Trp Leu Trp Ala Ser Asn Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp Val Val Lys Asp Trp Leu Ser Gln Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn Val Asp Ala Leu Leu Asn Trp Ala Tyr Ser Ser Gly Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Glu Ala Phe Asp Asn Lys Asn Ile Pro Ala Leu Val Tyr Ala Ile Gln Asn Gly Glu Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Asn Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile Leu Thr Tyr Glu Gly Gln Pro Val Ile Phe Tyr Arg Asp Tyr Glu Glu Trp Leu Asn Lys Asp Lys Leu Asn Asn Leu Ile Trp Ile His Glu His Leu Ala Gly Gly Ser Thr Lys Ile Leu Tyr Tyr Asp Asp Asp Glu Leu Ile Phe Met Arg Glu Gly Tyr Gly Asp Arg Pro Gly Leu Ile Thr Tyr Ile Asn Leu Gly Ser Asp Trp Ala Glu Arg Trp Val Asn Val Gly Ser Lys Phe Ala Gly Tyr Thr Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp Val Asp Arg Tyr Val Gln Tyr Asp Gly Trp Val Lys Leu Thr Ala Pro Pro His Asp Pro Ala Asn Gly Tyr Tyr Gly Tyr Ser Val Trp Ser Tyr Ala Gly Val Gly Arg Ser His His His His His His

SEQ ID NO: 117

ttgcgagtgttcctggttgtgccaaagctgagccgcccatttcaggcagagtcacaacaacaagacagggacataacaatgaaacacacagcg
ggaatgctggcgatcgcaggtatgctgatcgcccccttggcgcatgccgatgtcatactgcacgccttcaactggaaatacagtgaagtcaccg
ccaaggccgatctcatcaaggctgccggctacaagcaggtgctcatctcaccgcctctgaagtcctcgggcaacgagtggtgggctcgttacc
agccccaggatctgcgcctggtcgacacccccccttggcaacaagcaggatctggagcagctgatcgccgcgatgcagacccggggcattgc
cgtctacgcggacgtggtgctcaaccacatggccaacgaaagctggaagcgcagcgacctcaactaccccggcagcgagctgctgcaaag
ctacgccggcaatccggcctactttgaacgccagaagctctttggcgatctggggcagaacttcctcgccggccaggattttcatccggagggg
tgcatcaccgactggaacaatccgggccatgtccagtactggcgactgtgcggcggggcgggtgacaaggggctgccggatctggacccca
acaactgggtggtgaaccagcaacaggcttacctgcaggcgctcaaggggatggggatcaagggttttcgggtcgatgcggtcaagcacatg
agcgattaccagatcaacgccgtgttcaccccgagatcaaacaggggatgcacgtctttggcgaggtgatcaccacgggggcgccggca
acagcgactatgagaacttcctcaaaccctacctcgacagcagcggccaggggcctacgacttcccgctcttcgcctccctgcgtggagcgc
tgggctacggcggcagcatgaacctgctggccgatcccggtgcctatggtcaggcgctgccgggtagccgcgccgtcaccttcgccatcacc
cacgacatccccaccaacgacggtttccgctaccagatcctcaaccagaccgacgagagactggcctatgcctacctgctcggtcgcgatggc
ggttcgcctctggtctactccgatcacggtgaaaccagggacaaggacggattgcgctggcaggactactatctgcgcaccgatctcaaaggg
atgatccgcttccataacacagtgcagggtcaaccgatgcagctcatcggcagtaacgactgcttcgtgctgttcaagcgtggcaagcagggc
gtggtcggcatcaacaagtgcgactacgagcaggagtactggctcgataccgccagattcgagatgaactggtatcgcaactaccgggatgtg
ctcgaccagaatgccgtggtcaacgtgcagagccagtgggtaaggctgaccatcccggcccgcggcgccagaatgtggctgcaggagtga

SEQ ID NO: 118

Met Arg Val Phe Leu Val Val Pro Leu Leu Ser Arg Pro Phe Gln Ala Glu Ser Gln Gln Gln Asp Arg
Asp Ile Thr Met Lys His Thr Ala Gly Met Leu Ala Ile Ala Gly Met Leu Ile Ala Pro Leu Ala His
Ala Asp Val Ile Leu His Ala Phe Asn Trp Lys Tyr Ser Glu Val Thr Ala Lys Ala Asp Leu Ile Lys
Ala Ala Gly Tyr Lys Gln Val Leu Ile Ser Pro Pro Leu Lys Ser Ser Gly Asn Glu Trp Trp Ala Arg
Tyr Gln Pro Gln Asp Leu Arg Leu Val Asp Thr Pro Leu Gly Asn Lys Gln Asp Leu Glu Gln Leu Ile
Ala Ala Met Gln Thr Arg Gly Ile Ala Val Tyr Ala Asp Val Val Leu Asn His Met Ala Asn Glu Ser
Trp Lys Arg Ser Asp Leu Asn Tyr Pro Gly Ser Glu Leu Leu Gln Ser Tyr Ala Gly Asn Pro Ala Tyr
Phe Glu Arg Gln Lys Leu Phe Gly Asp Leu Gly Gln Asn Phe Leu Ala Gly Gln Asp Phe His Pro
Glu Gly Cys Ile Thr Asp Trp Asn Asn Pro Gly His Val Gln Tyr Trp Arg Leu Cys Gly Gly Ala Gly
Asp Lys Gly Leu Pro Asp Leu Asp Pro Asn Asn Trp Val Val Asn Gln Gln Gln Ala Tyr Leu Gln
Ala Leu Lys Gly Met Gly Ile Lys Gly Phe Arg Val Asp Ala Val Lys His Met Ser Asp Tyr Gln Ile
Asn Ala Val Phe Thr Pro Glu Ile Lys Gln Gly Met His Val Phe Gly Glu Val Ile Thr Thr Gly Gly
Ala Gly Asn Ser Asp Tyr Glu Asn Phe Leu Lys Pro Tyr Leu Asp Ser Ser Gly Gln Gly Ala Tyr Asp
Phe Pro Leu Phe Ala Ser Leu Arg Gly Ala Leu Gly Tyr Gly Gly Ser Met Asn Leu Leu Ala Asp Pro

FIGURE 16UU

Gly Ala Tyr Gly Gln Ala Leu Pro Gly Ser Arg Ala Val Thr Phe Ala Ile Thr His Asp Ile Pro Thr Asn Asp Gly Phe Arg Tyr Gln Ile Leu Asn Gln Thr Asp Glu Arg Leu Ala Tyr Ala Tyr Leu Leu Gly Arg Asp Gly Gly Ser Pro Leu Val Tyr Ser Asp His Gly Glu Thr Arg Asp Lys Asp Gly Leu Arg Trp Gln Asp Tyr Tyr Leu Arg Thr Asp Leu Lys Gly Met Ile Arg Phe His Asn Thr Val Gln Gly Gln Pro Met Gln Leu Ile Gly Ser Asn Asp Cys Phe Val Leu Phe Lys Arg Gly Lys Gln Gly Val Val Gly Ile Asn Lys Cys Asp Tyr Glu Gln Glu Tyr Trp Leu Asp Thr Ala Arg Phe Glu Met Asn Trp Tyr Arg Asn Tyr Arg Asp Val Leu Asp Gln Asn Ala Val Val Asn Val Gln Ser Gln Trp Val Arg Leu Thr Ile Pro Ala Arg Gly Ala Arg Met Trp Leu Gln Glu

SEQ ID NO: 119

atgcaaacgtttgcattcttattttactcaaagaaaggatgggtgtgcatgaattatttgaaaaaagtgtggttgtattacgctatcgtcgctaccttaa
tcatttcctttcttacacctttttcaacagcacaagctaatactgcacctgttaacggaacaatgatgcaatatttcgaatgggacttacctaatgatgg
gacgctttggacgaaagtaaaaaatgaagctaccaatctttcttcactaggtatcacagcactatggctccctccagcatataaaggaacgagcc
aaagcgatgtcggatacggtgtttacgatttatatgaccttggggaatttaatcaaaaagggacgatccgaacgaaatacggaacaaaaacaca
atatattcaagccattcaaactgcccaagccgcagggatgcaagtatatgcggatgttgtatttaatcataaggcaggggctgacagtacagaatt
tgtcgatgcagttgaggtaaacccttctaatcgaaatcaagaaacatctggcacatatcaaattcaagcatggacaaaatttgattttcctggtcgtg
gaaacacatactccagcttcaaatggcgctggtaccattttgatggtacggattgggacgaaagtcgtaaattaaatcgtatttacaaattccgcgg
tacaggaaaagcgtgggactgggaagtcgatacagaaaacggaaactatgattatttaatgttcgctgatttagatatggatcaccctgaggttgt
gacagaattaaaaaactggggaacgtggtacgtcaatactacaaatatcgatggattccgcttagatgccgtaaaacatattaaatacagcttttc
cctgactggctaacatatgtacgtaatcaaacaggaaaaaatttatttgccgttggggaattttggagctatgacgtcaataagctgcataattacat
tacaaaaacaaatgggtcgatgtcattatttgatgcacccttgcataacaacttttataccgcttccaaatcgagtggatattttgacatgcgttatttat
tgaataatacattaatgaaagatcaaccttcactcgctgtaacacttgtcgataaccacgacacgcaaccagggcaatctttacagtcatgggtcg
aaccttggtttaaacagcttgcttacgcctttatttttaacaagacaagaagggtatccttgcgtatttacggtgattattatggaatccctaaatacaat
atcccggggttaaaaagtaaaatcgacccgcttttaattgctcgtcgtgattacgcttatggaacacaacgtgattacattgatcatcaagacattat
cggatggacacgagaaggcattgatgcaaaaccgaactctggactggcggctttaattaccgacggtcctggtggaagtaaatggatgtatgtc
ggtaaaaagcatgccgggaaagtattttatgatttaactggaaatcgaagtgacacagtaacgattaatgcggatggttggggagaatttaaagta
aacggaggatccgtctcaatttgggtggctaaaacgtcaaacgtcacatttacagtcaataacgccacaacaacaagcggacaaaacgtatatg
ttgtcggcaacattccagagctaggcaattgtcgcacgggttaa

SEQ ID NO: 120

Met Gln Thr Phe Ala Phe Leu Phe Tyr Ser Lys Lys Gly Trp Val Cys Met Asn Tyr Leu Lys Lys Val Trp Leu Tyr Tyr Ala Ile Val Ala Thr Leu Ile Ile Ser Phe Leu Thr Pro Phe Ser Thr Ala Gln Ala Asn Thr Ala Pro Val Asn Gly Thr Met Met Gln Tyr Phe Glu Trp Asp Leu Pro Asn Asp Gly Thr Leu Trp Thr Lys Val Lys Asn Glu Ala Thr Asn Leu Ser Ser Leu Gly Ile Thr Ala Leu Trp Leu Pro Pro Ala Tyr Lys Gly Thr Ser Gln Ser Asp Val Gly Tyr Gly Val Tyr Asp Leu Tyr Asp Leu Gly Glu Phe Asn Gln Lys Gly Thr Ile Arg Thr Lys Tyr Gly Thr Lys Thr Gln Tyr Ile Gln Ala Ile Gln Thr Ala Glu Ala Ala Gly Met Gln Val Tyr Ala Asp Val Val Phe Asn His Lys Ala Gly Ala Asp Ser Thr Glu Phe Val Asp Ala Val Glu Val Asn Pro Ser Asn Arg Asn Gln Glu Thr Ser Gly Thr Tyr Gln Ile Gln Ala Trp Thr Lys Phe Asp Phe Pro Gly Arg Gly Asn Thr Tyr Ser Ser Phe Lys Trp Arg Trp Tyr His Phe Asp Gly Thr Asp Trp Asp Glu Ser Arg Lys Leu Asn Arg Ile Tyr Lys Phe Arg Gly Thr Gly Lys Ala Trp Asp Trp Glu Val Asp Thr Glu Asn Gly Asn Tyr Asp Tyr Leu Met Phe Ala Asp Leu Asp Met Asp His Pro Glu Val Val Thr Glu Leu Lys Asn Trp Gly Thr Trp Tyr Val Asn Thr Thr Asn Ile Asp Gly Phe Arg Leu Asp Ala Val Lys His Ile Lys Tyr Ser Phe Phe Pro Asp Trp Leu Thr Tyr Val Arg Asn Gln Thr Gly Lys Asn Leu Phe Ala Val Gly Glu Phe Trp Ser Tyr Asp Val Asn Lys Leu His Asn Tyr Ile Thr Lys Thr Asn Gly Ser Met Ser Leu Phe Asp Ala Pro Leu His Asn Asn Phe Tyr Thr Ala Ser Lys Ser Ser Gly Tyr Phe Asp Met Arg Tyr Leu Leu Asn Asn Thr Leu Met Lys Asp Gln Pro Ser Leu Ala Val Thr Leu Val Asp Asn His Asp Thr Gln Pro Gly Gln Ser Leu Gln Ser Trp Val Glu Pro Trp Phe Lys Gln Leu Ala Tyr Ala Phe Ile Leu Thr Arg Gln Glu Gly Tyr Pro Cys Val Phe Tyr Gly Asp Tyr Tyr Gly Ile Pro Lys Tyr Asn Ile Pro Gly Leu Lys Ser Lys Ile Asp Pro Leu Leu Ile Ala Arg Arg Asp Tyr Ala Tyr Gly Thr Gln Arg Asp Tyr Ile Asp His Gln Asp Ile Ile Gly Trp Thr

FIGURE 16VV

Arg Glu Gly Ile Asp Ala Lys Pro Asn Ser Gly Leu Ala Ala Leu Ile Thr Asp Gly Pro Gly Gly Ser Lys Trp Met Tyr Val Gly Lys Lys His Ala Gly Lys Val Phe Tyr Asp Leu Thr Gly Asn Arg Ser Asp Thr Val Thr Ile Asn Ala Asp Gly Trp Gly Glu Phe Lys Val Asn Gly Gly Ser Val Ser Ile Trp Val Ala Lys Thr Ser Asn Val Thr Phe Thr Val Asn Asn Ala Thr Thr Thr Ser Gly Gln Asn Val Tyr Val Val Gly Asn Ile Pro Glu Leu Gly Asn Cys Arg Thr Gly

SEQ ID NO: 121

atgctcgccctgtcgctcggcggctgcggcatcgacgcgggcccgacaggccctcgcgtcgtggagccgctgccgcagcgccccacgcttc
cgcaggagtaccgcgccagcggccacgcggccgccggcgacgtgttcgtgcacctgttcgagtggaagtggccggacatcgcggaggaat
gcgagaacgtgctggggccggcgggctacgaggcggtgcaggtgtcgccgccgcaggagcacctggtgcagcaggggcgccgtggtg
gcagcggtaccagccggtgagctactcggtggcgctgagccgcagcggcacgggcgtggagttcagcaacatgatcagccggtgcaaggc
cgccggcgtggacatctacgtggacgccgtcatcaaccacatgacggccggtgcggggacggggagcaacggcaccgcctacaccaagta
caactaccccggcctgtacgcgcaggcggactttcaccgcagtgcgcggtgggcgactacaccagcgccgccaacgtgcaggactgcga
actgctggggctggctgacctgaacaccggcgcggccggcgtgcagcagaagatcgcggactacctggtctcgctggcgcggctgggcgt
ggcgggttttcgcatcgacgccgccaagcacatccagccggtggaactggacgccatcgtggaccgcgtgaaccagacgctggcggcgga
ggggcgcccgcttccctactggttcgccgaggtgatcgacaacggcggcgaggggtgcggcgcgagcactactacggcctgggatacgg
caccggcggcgccgcggacatcacggagttccgctacaagggcgtgggcgacaagttcctgggcagcggcggccagcggctggtggacc
tgaagaacttctcggcggtgacgtggaacctgatgccgtcggacaaggccgtcgtctttctggagaaccacgatacgcagcgcggcggcgc
atcggctaccgcgatggcacggcgttccggctggccaacgtgtggatgctggcgcagccgtacggctatccgtcggtgatgtccagctacgc
ctttgaccgcacctccccctttggccgcgacgccggcccgccctccgaggacggcgcgacgaaggacgtgacgtgcgcgcccacgctgga
gacggcggtgctgggcacctgggtgtgcgagcaccgcgaccccgtcattcagcggatggtgggctttcgccgcgcgatggcgggcacgga
cctgaaccgctggtgggacaacggcggcaacgccattgccttttcgcgcggggaccggggcttcgtcgccatcagccgcgagccgaaggtg
accatggcggccgtgcccagcggactgtcccccggcacctactgcgacgtgctgaccggcggcaaggtgggcaacgcctgcgcgggaac
cagcgtgacggtcgactctcagggcgtggtgcagctgagcatcgtcgagaactcggctctggtgatccacctcggggccaagctgtaacggc
gcgctggcggatgtgcggaggg

SEQ ID NO: 122

Met Leu Ala Leu Ser Leu Gly Gly Cys Gly Ile Asp Ala Gly Pro Thr Gly Pro Arg Val Val Glu Pro Leu Pro Gln Arg Pro Thr Leu Pro Gln Glu Tyr Arg Ala Ser Gly His Ala Ala Ala Gly Asp Val Phe Val His Leu Phe Glu Trp Lys Trp Pro Asp Ile Ala Glu Glu Cys Glu Asn Val Leu Gly Pro Ala Gly Tyr Glu Ala Val Gln Val Ser Pro Pro Gln Glu His Leu Val Gln Gln Gly Ala Pro Trp Trp Gln Arg Tyr Gln Pro Val Ser Tyr Ser Val Ala Leu Ser Arg Ser Gly Thr Gly Val Glu Phe Ser Asn Met Ile Ser Arg Cys Lys Ala Ala Gly Val Asp Ile Tyr Val Asp Ala Val Ile Asn His Met Thr Ala Gly Ala Gly Thr Gly Ser Asn Gly Thr Ala Tyr Thr Lys Tyr Asn Tyr Pro Gly Leu Tyr Ala Gln Ala Asp Phe His Pro Gln Cys Ala Val Gly Asp Tyr Thr Ser Ala Ala Asn Val Gln Asp Cys Glu Leu Leu Gly Leu Ala Asp Leu Asn Thr Gly Ala Ala Gly Val Gln Gln Lys Ile Ala Asp Tyr Leu Val Ser Leu Ala Arg Leu Gly Val Ala Gly Phe Arg Ile Asp Ala Ala Lys His Ile Gln Pro Val Glu Leu Asp Ala Ile Val Asp Arg Val Asn Glu Thr Leu Ala Ala Glu Gly Arg Pro Leu Pro Tyr Trp Phe Ala Glu Val Ile Asp Asn Gly Gly Glu Gly Val Arg Arg Glu His Tyr Tyr Gly Leu Gly Tyr Gly Thr Gly Gly Ala Ala Asp Ile Thr Glu Phe Arg Tyr Lys Gly Val Gly Asp Lys Phe Leu Gly Ser Gly Gly Gln Arg Leu Val Asp Leu Lys Asn Phe Ser Ala Val Thr Trp Asn Leu Met Pro Ser Asp Lys Ala Val Val Phe Leu Glu Asn His Asp Thr Gln Arg Gly Gly Gly Ile Gly Tyr Arg Asp Gly Thr Ala Phe Arg Leu Ala Asn Val Trp Met Leu Ala Gln Pro Tyr Gly Tyr Pro Ser Val Met Ser Ser Tyr Ala Phe Asp Arg Thr Ser Pro Phe Gly Arg Asp Ala Gly Pro Pro Ser Glu Asp Gly Ala Thr Lys Asp Val Thr Cys Ala Pro Thr Leu Glu Thr Ala Val Leu Gly Thr Trp Val Cys Glu His Arg Asp Pro Val Ile Gln Arg Met Val Gly Phe Arg Arg Ala Met Ala Gly Thr Asp Leu Asn Arg Trp Trp Asp Asn Gly Gly Asn Ala Ile Ala Phe Ser Arg Gly Asp Arg Gly Phe Val Ala Ile Ser Arg Glu Pro Lys Val Thr Met Ala Ala Val Pro Ser Gly Leu Ser Pro Gly Thr Tyr Cys Asp Val Leu Thr Gly Gly Lys Val Gly Asn Ala Cys Ala Gly Thr Ser Val Thr Val Asp Ser Gln Gly Val Val Gln Leu Ser Ile Val Glu Asn Ser Ala Leu Val Ile His Leu Gly Ala Lys Leu Arg Arg Ala Gly Gly Cys Ala Glu

FIGURE 16WW

SEQ ID NO: 123

atgccccaggccattcgcactttttcacgttggacgttgttcggcttaatcggcgttttctgcttggtctcgtcttttctgtcccaccccgggcaatcc
aggcccagacaaccccggcccgtaccgttatggttcacctcttcgagtggaaatggaccgacatcgctaaagaatgcgagaatttcctcggac
cgaaaggctttgccgcaatccaggtatcgccgccccaggagcatgtccaggggtcgcaatggtggacccgctatcagccggtcagctacaag
atcgagagccgctccggcacccgggccgagttcgccaatatggtctcgcgctgcaaagccgtcggggtcgatatctatgtcgatgccgtgatc
aaccatatgacgactgtcggctccggcactggtatggctggatcgacctacaccagctacacctatccggggctgtatcagacccaggacttcc
accactgcgggcgcaatggcaacgatgatatcagcagctacggcgatcgctgggaagtacaaaactgcgaactgctcaacctagccgacctc
aacaccggcgctgagtatgtccggggtaaactcgccgcctatatgaacgatctgcgcggcctgggcgtcgccggatttcggatcgatgccgcc
aagcacatggataccaacgacatcaacaatatcgttggccgcctgcccaacgcgccctacatctaccaggaagtgatcgaccagggcggcga
gccaattaccgccggcgaatacttccagaatggcgatgtgaccgagttcaagtacagccgcgagatctcgcgcatgttcaaaaccggccagct
gacccatatgagccagttcggcactgcctggggcttcatgtccagcgacctggcagtagttttcaccgataaccacgacaaccagcgcggtca
cggcggcgccggcgatgtcttgacctacaaagatggccagctgtacaccctgggcaatatcttcgagctagcctggccgtatggctacccaca
ggtcatgtcgagctacacgttcagcaacggcgaccagggccgccatcgaccaatgtgtacgcaaccacaacgcctgattgtggcaacggcc
gctgggtctgtgagcaccgctggcgaggaatcgccaacatggtcgcgttccgcaactacaccgccccgaccttcagcaccagcaactggtgg
agcaacggcaacaaccagatcgctttcagccgcgggaccctgggctttgtggcgatcaatcgggaaggtggcagcctgaaccgcaccttcca
aaccggcctgcccgtcggcacctactgcgatgtcattcacggcgatttcaatgccagcgccggcacctgttccggcccaactatcgctgtcaac
ggctccggacaggcaaccatcacggtcaacgcgatggacgcggtggcgatctacggcggagccaggctcgccactccggccagtgtcaac
gtgacattcaacgaaaacgccacgaccacctgggggcagaatgtgtatatcgtcggcaacgtcgccgccctgggcagctggaacgcaggca
gcgcggtcttactctcctccgctaactacccaatctggagcaagaccatcgccctgccagccaacaccgccattgagtacaagtacatcaaaaa
ggatggcgcgggcaatgtggtgtgggaaagcggcgccaaccgcgtctttaccaccccggcagcggcagtgccacgcgcaacgatacctg
gaaatag

SEQ ID NO: 124

Met Pro Gln Ala Ile Arg Thr Phe Ser Arg Trp Thr Leu Phe Gly Leu Ile Gly Val Phe Leu Leu Gly
Leu Val Phe Ser Val Pro Pro Arg Ala Ile Gln Ala Gln Thr Thr Pro Ala Arg Thr Val Met Val His
Leu Phe Glu Trp Lys Trp Thr Asp Ile Ala Lys Glu Cys Glu Asn Phe Leu Gly Pro Lys Gly Phe Ala
Ala Ile Gln Val Ser Pro Pro Gln Glu His Val Gln Gly Ser Gln Trp Trp Thr Arg Tyr Gln Pro Val Ser
Tyr Lys Ile Glu Ser Arg Ser Gly Thr Arg Ala Glu Phe Ala Asn Met Val Ser Arg Cys Lys Ala Val
Gly Val Asp Ile Tyr Val Asp Ala Val Ile Asn His Met Thr Thr Val Gly Ser Gly Thr Gly Met Ala
Gly Ser Thr Tyr Thr Ser Tyr Thr Tyr Pro Gly Leu Tyr Gln Thr Gln Asp Phe His His Cys Gly Arg
Asn Gly Asn Asp Asp Ile Ser Ser Tyr Gly Asp Arg Trp Glu Val Gln Asn Cys Glu Leu Leu Asn
Leu Ala Asp Leu Asn Thr Gly Ala Glu Tyr Val Arg Gly Lys Leu Ala Ala Tyr Met Asn Asp Leu
Arg Gly Leu Gly Val Ala Gly Phe Arg Ile Asp Ala Ala Lys His Met Asp Thr Asn Asp Ile Asn Asn
Ile Val Gly Arg Leu Pro Asn Ala Pro Tyr Ile Tyr Gln Glu Val Ile Asp Gln Gly Gly Glu Pro Ile Thr
Ala Gly Glu Tyr Phe Gln Asn Gly Asp Val Thr Glu Phe Lys Tyr Ser Arg Glu Ile Ser Arg Met Phe
Lys Thr Gly Gln Leu Thr His Met Ser Gln Phe Gly Thr Ala Trp Gly Phe Met Ser Ser Asp Leu Ala
Val Val Phe Thr Asp Asn His Asp Asn Gln Arg Gly His Gly Gly Ala Gly Asp Val Leu Thr Tyr
Lys Asp Gly Gln Leu Tyr Thr Leu Gly Asn Ile Phe Glu Leu Ala Trp Pro Tyr Gly Tyr Pro Gln Val
Met Ser Ser Tyr Thr Phe Ser Asn Gly Asp Gln Gly Pro Pro Ser Thr Asn Val Tyr Ala Thr Thr Thr
Pro Asp Cys Gly Asn Gly Arg Trp Val Cys Glu His Arg Trp Arg Gly Ile Ala Asn Met Val Ala Phe
Arg Asn Tyr Thr Ala Pro Thr Phe Ser Thr Ser Asn Trp Trp Ser Asn Gly Asn Asn Gln Ile Ala Phe
Ser Arg Gly Thr Leu Gly Phe Val Ala Ile Asn Arg Glu Gly Gly Ser Leu Asn Arg Thr Phe Gln Thr
Gly Leu Pro Val Gly Thr Tyr Cys Asp Val Ile His Gly Asp Phe Asn Ala Ser Ala Gly Thr Cys Ser
Gly Pro Thr Ile Ala Val Asn Gly Ser Gly Gln Ala Thr Ile Thr Val Asn Ala Met Asp Ala Val Ala Ile
Tyr Gly Gly Ala Arg Leu Ala Thr Pro Ala Ser Val Asn Val Thr Phe Asn Glu Asn Ala Thr Thr Thr
Trp Gly Gln Asn Val Tyr Ile Val Gly Asn Val Ala Ala Leu Gly Ser Trp Asn Ala Gly Ser Ala Val
Leu Leu Ser Ser Ala Asn Tyr Pro Ile Trp Ser Lys Thr Ile Ala Leu Pro Ala Asn Thr Ala Ile Glu Tyr
Lys Tyr Ile Lys Lys Asp Gly Ala Gly Asn Val Val Trp Glu Ser Gly Ala Asn Arg Val Phe Thr Thr
Pro Gly Ser Gly Ser Ala Thr Arg Asn Asp Thr Trp Lys

FIGURE 16XX

SEQ ID NO: 125

gtggtgcacatgaagttgaagtaccttgccttagttttgttggctgtggcttcgataggcctactctcgactccagtgggtgctgccaagtactccg
aactcgaagagggcggtgttataatgcaggccttctactgggatgttcccggaggggaatctggtgggacaccataagacagaaaatcccg
gagtggtacgacgctggaatctcggcgatatggattcctccagctagcaaagggatgggcggtggttattccatgggctacgatccctacgattt
ctttgacctcggcgagtactatcagaagggaacagttgagacgcgcttcggctcaaaggaggaactggtgaacatgataaacaccgcacactc
ctatggcataaaggtgatagcggacatagtcataaaccaccgcgccggtggagaccttgagtggaaccccctttgtaaacaactatacttggaca
gacttctccaaggtcgcctccggtaaatacacggccaactaccttgacttccacccaaacgaggtcaagtgctgcgatgagggtacatttggtga
ctttccggacatcgcccacgagaagagctgggatcagtactggctctgggcaagcaatgagagctacgccgcatatctccggagcataggga
tcgatgcatggcgtttcgactacgtcaaaggttacggagcgtgggttgttaatgactggctcagctggtggggaggctgggccgttggagagta
ctgggacacgaacgttgatgcactccttaactgggcatacgacagcggtgccaaggtctttgacttcccgctctactacaagatggacgaagcc
tttgacaacaccaacatcccgctttggtttacgccctccagaacggaggaacagtcgtttcccgcgatcccttcaaggcagtaactttcgttgcc
aaccacgatacagatataatctggaacaagtatccggcttatgcgttcatccttacctatgagggacagcctgttatattttaccgcgactacgagg
agtggctcaacaaggataagcttaacaaccttatctggatacacgagcaccttgccggaggaagtaccaagatcctctactacgataacgatga
gctaatattcatgagggagggctacgggagcaagccgggcctcataacctacataaacctcggaaacgactgggccgagcgctgggtgaac
gtcggctcaaagtttgccggctacacaatccatgaatacacaggcaatctcggtggctgggttgacaggtgggttcagtacgatggatgggtta
aactgacggcacctcctcatgatccagccaacggatattacggctactcagtctggagctacgcaggcgtcggatga

SEQ ID NO: 126

Val Val His Met Lys Leu Lys Tyr Leu Ala Leu Val Leu Leu Ala Val Ala Ser Ile Gly Leu Leu Ser
Thr Pro Val Gly Ala Ala Lys Tyr Ser Glu Leu Glu Glu Gly Gly Val Ile Met Gln Ala Phe Tyr Trp
Asp Val Pro Gly Gly Gly Ile Trp Trp Asp Thr Ile Arg Gln Lys Ile Pro Glu Trp Tyr Asp Ala Gly Ile
Ser Ala Ile Trp Ile Pro Pro Ala Ser Lys Gly Met Gly Gly Gly Tyr Ser Met Gly Tyr Asp Pro Tyr
Asp Phe Phe Asp Leu Gly Glu Tyr Tyr Gln Lys Gly Thr Val Glu Thr Arg Phe Gly Ser Lys Glu Glu
Leu Val Asn Met Ile Asn Thr Ala His Ser Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His Arg
Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Asn Asn Tyr Thr Trp Thr Asp Phe Ser Lys Val Ala
Ser Gly Lys Tyr Thr Ala Asn Tyr Leu Asp Phe His Pro Asn Glu Val Lys Cys Cys Asp Glu Gly Thr
Phe Gly Asp Phe Pro Asp Ile Ala His Glu Lys Ser Trp Asp Gln Tyr Trp Leu Trp Ala Ser Asn Glu
Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly
Ala Trp Val Val Asn Asp Trp Leu Ser Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn
Val Asp Ala Leu Leu Asn Trp Ala Tyr Asp Ser Gly Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys
Met Asp Glu Ala Phe Asp Asn Thr Asn Ile Pro Ala Leu Val Tyr Ala Leu Gln Asn Gly Gly Thr Val
Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Asp Ile Ile Trp Asn Lys
Tyr Pro Ala Tyr Ala Phe Ile Leu Thr Tyr Glu Gly Gln Pro Val Ile Phe Tyr Arg Asp Tyr Glu Glu
Trp Leu Asn Lys Asp Lys Leu Asn Asn Leu Ile Trp Ile His Glu His Leu Ala Gly Gly Ser Thr Lys
Ile Leu Tyr Tyr Asp Asn Asp Glu Leu Ile Phe Met Arg Glu Gly Tyr Gly Ser Lys Pro Gly Leu Ile
Thr Tyr Ile Asn Leu Gly Asn Asp Trp Ala Glu Arg Trp Val Asn Val Gly Ser Lys Phe Ala Gly Tyr
Thr Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp Val Asp Arg Trp Val Gln Tyr Asp Gly Trp Val
Lys Leu Thr Ala Pro Pro His Asp Pro Ala Asn Gly Tyr Tyr Gly Tyr Ser Val Trp Ser Tyr Ala Gly
Val Gly

SEQ ID NO: 127

gtgtgcatgaattatttgaaaaaagtgtggttgtattacgctatcgtcgctaccttaatcatttacttttcttacgccctttttcaactgcacaagccaacac
tgcaccagtcaacggaacgatgatgcaatatttcgaatgggatttaccgaatgatggcacactttggacgaaagtaaaaaacgaagcaagcagt
ctttcttctttaggtattactgcgttatggttaccacctgcatacaaaggaacgagccaaggggatgtcgggtatggcgtgtacgatttgtatgactt
aggagaatttaatcaaaaagggacgattcgaacgaaatacggaacaaaaacgcaatatttacaagccattcaagcggcaaaaagcgctggcat
gcaagtatacgctgatgtcgtatttaatcacaaggcgggggcagatagtacagaatgggttgacgcagtcgaagtgaatccttctaatcgaaacc
aagaaacatctggcacatatcaaattcaagcatggacaaaatttgatttccctggccgtgggaacacatactcaagctttaaatggcgatggtatc
attttgacggtacggattgggatgaaagccgaaaactaaatcgtatttacaaatttcgtggcacaggaaaagcatgggattgggaagtagacaca
gagaacggaaactatgactacttaatgtttgctgatttagatatggatcaccctgaagtcgtgacagagctaaaaaactggggaacatggtacgtc

FIGURE 16YY

aatacgacaaatgtcgatgggtttcgcttagatgcagtaaagcatattaaatatagcttcttcccagattggttaacacatgtgcgttcacaaacacg
aaaaaatctttttgcagtaggagaattttggagctacgatgtcaataaactgcataactacattacaaaaacaagtggaaccatgtcgttatttgatg
cgccacttcataacaacttttacactgcttcaaaatctagcgggtattttgacatgcgctatttgttaaataatacgttgatgaaagaccagccttctct
tgcggtcacactcgttgataatcatgacacgcaaccgggacaatctttacaatcatgggtagagccttggtttaagccgcttgcttatgccttattttt
gacaagacaagaaggatatccttgcgtattttacggcgactattacggcatccctaaatacaacattccgggattgaaaagtaaaatcgatccgct
tctcattgcccgtagagactacgcatacggaacacaacgtgattatattgaccatcaagacattattggatggacacgggaaggaattgactcaa
aaccgaactctggacttgcggctttaattactgacggccctggtggaagtaaatggatgtatgtaggtaaaaagcatgctggaaaagtgttttacg
atctcactggaaatcgaagcgatacggtaacgattaatgcagacggctggggagagtttaaagtaaacggtggctccgtttccatttgggttgcc
aaaacatcacaagtcacgtttaccgtcaacaatgcgacaacgataagcggacaaaatgtgtatgtcgttggtaacattccagagctcggaaattg
gaacacagcaaacgcaatcaaaatgaccccatcttcttatccaacgtggaaagcaaccattgctcttccacaaggaaaagccattgaatttaaatt
tattaaaaaagaccaatcgggaaatgttgtttgggaaagcattccaaaccgaacatacaccgttccatttttatcaacaggctcatatacagctagtt
ggaatgtaccttaa

SEQ ID NO: 128
Val Cys Met Asn Tyr Leu Lys Lys Val Trp Leu Tyr Tyr Ala Ile Val Ala Thr Leu Ile Ile Tyr Phe
Leu Thr Pro Phe Ser Thr Ala Gln Ala Asn Thr Ala Pro Val Asn Gly Thr Met Met Gln Tyr Phe Glu
Trp Asp Leu Pro Asn Asp Gly Thr Leu Trp Thr Lys Val Lys Asn Glu Ala Ser Ser Leu Ser Ser Leu
Gly Ile Thr Ala Leu Trp Leu Pro Pro Ala Tyr Lys Gly Thr Ser Gln Gly Asp Val Gly Tyr Gly Val
Tyr Asp Leu Tyr Asp Leu Gly Glu Phe Asn Gln Lys Gly Thr Ile Arg Thr Lys Tyr Gly Thr Lys Thr
Gln Tyr Leu Gln Ala Ile Gln Ala Ala Lys Ser Ala Gly Met Gln Val Tyr Ala Asp Val Val Phe Asn
His Lys Ala Gly Ala Asp Ser Thr Glu Trp Val Asp Ala Val Glu Val Asn Pro Ser Asn Arg Asn Gln
Glu Thr Ser Gly Thr Tyr Gln Ile Gln Ala Trp Thr Lys Phe Asp Phe Pro Gly Arg Gly Asn Thr Tyr
Ser Ser Phe Lys Trp Arg Trp Tyr His Phe Asp Gly Thr Asp Trp Asp Glu Ser Arg Lys Leu Asn Arg
Ile Tyr Lys Phe Arg Gly Thr Gly Lys Ala Trp Asp Trp Glu Val Asp Thr Glu Asn Gly Asn Tyr Asp
Tyr Leu Met Phe Ala Asp Leu Asp Met Asp His Pro Glu Val Val Thr Glu Leu Lys Asn Trp Gly
Thr Trp Tyr Val Asn Thr Thr Asn Val Asp Gly Phe Arg Leu Asp Ala Val Lys His Ile Lys Tyr Ser
Phe Phe Pro Asp Trp Leu Thr His Val Arg Ser Gln Thr Arg Lys Asn Leu Phe Ala Val Gly Glu Phe
Trp Ser Tyr Asp Val Asn Lys Leu His Asn Tyr Ile Thr Lys Thr Ser Gly Thr Met Ser Leu Phe Asp
Ala Pro Leu His Asn Asn Phe Tyr Thr Ala Ser Lys Ser Ser Gly Tyr Phe Asp Met Arg Tyr Leu Leu
Asn Asn Thr Leu Met Lys Asp Gln Pro Ser Leu Ala Val Thr Leu Val Asp Asn His Asp Thr Gln
Pro Gly Gln Ser Leu Gln Ser Trp Val Glu Pro Trp Phe Lys Pro Leu Ala Tyr Ala Phe Ile Leu Thr
Arg Gln Glu Gly Tyr Pro Cys Val Phe Tyr Gly Asp Tyr Tyr Gly Ile Pro Lys Tyr Asn Ile Pro Gly
Leu Lys Ser Lys Ile Asp Pro Leu Leu Ile Ala Arg Arg Asp Tyr Ala Tyr Gly Thr Gln Arg Asp Tyr
Ile Asp His Gln Asp Ile Ile Gly Trp Thr Arg Glu Gly Ile Asp Ser Lys Pro Asn Ser Gly Leu Ala Ala
Leu Ile Thr Asp Gly Pro Gly Gly Ser Lys Trp Met Tyr Val Gly Lys Lys His Ala Gly Lys Val Phe
Tyr Asp Leu Thr Gly Asn Arg Ser Asp Thr Val Thr Ile Asn Ala Asp Gly Trp Gly Glu Phe Lys Val
Asn Gly Gly Ser Val Ser Ile Trp Val Ala Lys Thr Ser Gln Val Thr Phe Thr Val Asn Asn Ala Thr
Thr Ile Ser Gly Gln Asn Val Tyr Val Val Gly Asn Ile Pro Glu Leu Gly Asn Trp Asn Thr Ala Asn
Ala Ile Lys Met Thr Pro Ser Ser Tyr Pro Thr Trp Lys Ala Thr Ile Ala Leu Pro Gln Gly Lys Ala Ile
Glu Phe Lys Phe Ile Lys Lys Asp Gln Ser Gly Asn Val Val Trp Glu Ser Ile Pro Asn Arg Thr Tyr
Thr Val Pro Phe Leu Ser Thr Gly Ser Tyr Thr Ala Ser Trp Asn Val Pro

SEQ ID NO: 129
ttgcgttgccgccgtggcagggacgggtgttggtgcgggcggcgtaatgcgctgccgcgacacccgcgtgaacaaaataatatgaattatttg
aataggatgggggtgtcaagaatgacaaaatctcgagagttgcggtgttcatggaaagtatttgttgttgggtgcctgttgtggatggcttgggga
tcttccgcgtccgccggcgtattgatgcaaggcttctactgggacgccagtaccgggaccagtgattcgtggtggacgcatttggccaagcaag
ccaacggtctaaaacgggcggggttcaccgccgtatggattcctccggtgcttaaaggggcttcaggggctattccaacgggtacgatcccctt
tgacgactatgatatcggaagcaaggaccagaaaggtaccgtggcgacgcgatgggggacgcgagaagaactgcaacgtgccgtggccgt
gatgcgcgcgaacggtctggatgtgtatgtggatctggtgctgaaccaccgcaacggggacgacgggaattggaattttcattacaaagatgc
gtacggcaaagtgggttacgggcggtttcaaaaggggttttacgattttcaccccaactacaacattcaggatgccaatgttcccaacgaggattc

FIGURE 16ZZ

cagcttcgggcgcgatttagcccatgacaatccgtatgtggccgatggactgaaggctgcaggcgattggctgaccaaagccctcgatgttca
gggatatcgtctggattacgtgaaaggcatcagctacaccttcctgaaaagttatctgtcctatggggccatgaacggaaaatttgccgtcggtga
gtactgggatgccaaccgggatacgttgaactggtgggcgaacacggcgatggaagggcgggcccatgtgtttgattttgcgttgcgcgagg
agctgaaaaacatgtgcaatgcggacgggtactacgacatgcgtcgattggaccacgcgggtctggtcggaatcgacccgtggaaggcggt
gacgtttgtcgaaaaccatgatacggatcggcacgaccccatctacaataacaagcatttggcgtatgcctacatcttgacgtcggaagggtatc
cgacggtgttctggaaggattactaccaatacggaatgaagccgatcatcgacaacctcatttggatccacgaacacattgcgtacggaacgac
ccaagagcgttggaaagacgaagatgtctttgtgtatgagcggaccggaggcaagcggctattggtggggcttaacgacaatcgcgccacca
gcaaaacggtcaccgtacagaccggctttggtgccaacgtggccttgcacgactacacggcaacggccccgatctccgtaccgacgcctac
ggtcgggtaaccttgaccattcctgcaaacgggtacgtggcctattccgttccgggcatctccggatcctttgtgccggtcgagaaaaccgtgac
gcaggagtttgccggggcgtccgacttggatattcgtccggccgataacacgcaatttgtgcaggtcgggcggatatacgccaaggcaaacaa
gccggttacagcggaattgtattgggatgccaaagactggacgacctccacgtcgattctcctagaagtgcgttcggcttcgggaacgctcatc
acgacaaagaccgtgacccaattgtcgtcccagggtacccgcgtttccttcacgccttcggctaccggatggtacgtcttttccattcgaagctat
aacacgccttcgacgaacccaaagccggcctactggttaaaggtaacgtatacggcgccgcaattgcttcagtaa

SEQ ID NO: 130

Met Arg Cys Arg Arg Gly Arg Asp Gly Cys Trp Cys Gly Arg Arg Asn Ala Leu Pro Arg His Pro Arg Glu Gln Asn Asn Met Asn Tyr Leu Asn Arg Met Gly Val Ser Arg Met Thr Lys Ser Arg Glu Leu Arg Cys Ser Trp Lys Val Phe Val Val Gly Cys Leu Leu Trp Met Ala Trp Gly Ser Ser Ala Ser Ala Gly Val Leu Met Gln Gly Phe Tyr Trp Asp Ala Ser Thr Gly Thr Ser Asp Ser Trp Trp Thr His Leu Ala Lys Gln Ala Asn Gly Leu Lys Arg Ala Gly Phe Thr Ala Val Trp Ile Pro Pro Val Leu Lys Gly Ala Ser Gly Gly Tyr Ser Asn Gly Tyr Asp Pro Phe Asp Asp Tyr Asp Ile Gly Ser Lys Asp Gln Lys Gly Thr Val Ala Thr Arg Trp Gly Thr Arg Glu Glu Leu Gln Arg Ala Val Ala Val Met Arg Ala Asn Gly Leu Asp Val Tyr Val Asp Leu Val Leu Asn His Arg Asn Gly Asp Asp Gly Asn Trp Asn Phe His Tyr Lys Asp Ala Tyr Gly Lys Val Gly Tyr Gly Arg Phe Gln Lys Gly Phe Tyr Asp Phe His Pro Asn Tyr Asn Ile Gln Asp Ala Asn Val Pro Asn Glu Asp Ser Ser Phe Gly Arg Asp Leu Ala His Asp Asn Pro Tyr Val Ala Asp Gly Leu Lys Ala Ala Gly Asp Trp Leu Thr Lys Ala Leu Asp Val Gln Gly Tyr Arg Leu Asp Tyr Val Lys Gly Ile Ser Tyr Thr Phe Leu Lys Ser Tyr Leu Ser Tyr Gly Ala Met Asn Gly Lys Phe Ala Val Gly Glu Tyr Trp Asp Ala Asn Arg Asp Thr Leu Asn Trp Trp Ala Asn Thr Ala Met Glu Gly Arg Ala His Val Phe Asp Phe Ala Leu Arg Glu Glu Leu Lys Asn Met Cys Asn Ala Asp Gly Tyr Tyr Asp Met Arg Arg Leu Asp His Ala Gly Leu Val Gly Ile Asp Pro Trp Lys Ala Val Thr Phe Val Glu Asn His Asp Thr Asp Arg His Asp Pro Ile Tyr Asn Asn Lys His Leu Ala Tyr Ala Tyr Ile Leu Thr Ser Glu Gly Tyr Pro Thr Val Phe Trp Lys Asp Tyr Tyr Gln Tyr Gly Met Lys Pro Ile Ile Asp Asn Leu Ile Trp Ile His Glu His Ile Ala Tyr Gly Thr Thr Gln Glu Arg Trp Lys Asp Glu Asp Val Phe Val Tyr Glu Arg Thr Gly Gly Lys Arg Leu Leu Val Gly Leu Asn Asp Asn Arg Ala Thr Ser Lys Thr Val Thr Val Gln Thr Gly Phe Gly Ala Asn Val Ala Leu His Asp Tyr Thr Gly Asn Gly Pro Asp Leu Arg Thr Asp Ala Tyr Gly Arg Val Thr Leu Thr Ile Pro Ala Asn Gly Tyr Val Ala Tyr Ser Val Pro Gly Ile Ser Gly Ser Phe Val Pro Val Glu Lys Thr Val Thr Gln Glu Phe Ala Gly Ala Ser Asp Leu Asp Ile Arg Pro Ala Asp Asn Thr Gln Phe Val Gln Val Gly Arg Ile Tyr Ala Lys Ala Asn Lys Pro Val Thr Ala Glu Leu Tyr Trp Asp Ala Lys Asp Trp Thr Thr Ser Thr Ser Ile Leu Leu Glu Val Arg Ser Ala Ser Gly Thr Leu Ile Thr Thr Lys Thr Val Thr Gln Leu Ser Ser Gln Gly Thr Arg Val Ser Phe Thr Pro Ser Ala Thr Gly Trp Tyr Val Phe Ser Ile Arg Ser Tyr Asn Thr Pro Ser Thr Asn Pro Lys Pro Ala Tyr Trp Leu Lys Val Thr Tyr Thr Ala Pro Gln Leu Leu Gln

SEQ ID NO: 131

atgccgcagctttacccattgccgccgcgctggcggcgcggcccggcagggcctggccgccttgacgctggccaccacggccctgggc
atctcgacggcccaggcccagagtgcaccgcgcacggccttcgtgcatctgttcgaatggaagtggaccgacatcgcgcgcgagtgcgaga
ccttcctcgggcccaagggcttcgcggcggtgcaggtgtcgccccgaacgagcacaactgggtgaccagcggtgatggtgcaccttatccg
tggtggatgcgctaccagccggtgagctacagcctggaccgcagccgcagcggcacgcgcgccgagttccaggacatggtcaaccgatgc
aatgccgtgggcgtgggcatctacgtggacgccgtgatcaatcacatgtccggcggcacgggcggcacctcgagcgctgggcgcagctgg

FIGURE 16AAA

agctatcacaactaccctgggctctatggccccaacgacttccaccagccggtgtgcagcatcaccaactacggggatgcgaacaatgtgcag
cgttgcgagctctcgggcttgcaggacctggacactgggagcgcttatgtgcgcggcaagatcgccgactatctggtggatctggtcaacatg
ggggtcaagggcttccgggtggatgcggccaagcacatcagcccgaccgacctgggcgccatcatcgatgcggtcaacagccgcaccggc
gcgaaccgccctttctggtttctggaggtgattggcgcggccggcgaggcagtgcagccgaaccagtacttctcgctcggcggcggccaggt
caccgtgaccgagttcaactatgggaagcaaatcttcggcaagttcgccggtggcggccgtctggccgagctgcgcagcttcggtgaaacctg
gggcctgatgcccagcagcaaagcgattgctttcatcgacaaccacgacaagcagcgcggtcatggcggcggtggcaactatctgacctacc
accatggctcgacctacgatctggccaacatcttcatgctggcttggccttatggctacccggcgctgatgtccagctatgccttcaaccgcagc
acggcctacgacacgagctttggcccgccacacgacagtggtggcgccacccgtggcccctgggatggtggcggcagccagccggcctgc
ttcaaccagagcatcggtggctgggtgtgtgagcaccgctggcggggcatcgccaatatggtggccttccgcaacgccacgctgcccaactg
gaccgtgaccgactggtgggacaacggcaacaaccagatcgctttcgggcggggtgacaagggcttcgtggtgatcaaccgcgaagacgc
cgcgctgacgcgcaacttcaagaccagcctgccagccggccagtactgcgatgtcatctccggggacttcaacaatggtcagtgcacgggcc
atgtggtgacggtcgatgccggcggctacgtgacgctgacggccgggcccaatggtgcggcggccatccacgtgggcgcccgtctggacg
gcgcctctcagccgccgacgaccgcctcggtgacgttcaacgcgtcggccgataccttttggggacagaacctgttcgtcgtgggcaaccaca
gcgcactgggcaactggtcgccggcggccgccaggccgatgacttggatttccggttcgggcacgcgcgggaactggcgcgcggtgctca
atttgccggccaataccacctaccaatacaagttcatcaagaaggacggggctggaaacgtggtttgggagggcggtggcaatcgcgtcgtga
ccacgccgtctgggggcggatcggtgagcacgggcggcaattggcagtag

SEQ ID NO: 132
Met Pro Gln Leu Tyr Pro Leu Pro Pro Arg Trp Arg Arg Ala Ala Arg Gln Gly Leu Ala Ala Leu Thr
Leu Ala Thr Thr Ala Leu Gly Ile Ser Thr Ala Gln Ala Gln Ser Ala Pro Arg Thr Ala Phe Val His
Leu Phe Glu Trp Lys Trp Thr Asp Ile Ala Arg Glu Cys Glu Thr Phe Leu Gly Pro Lys Gly Phe Ala
Ala Val Gln Val Ser Pro Pro Asn Glu His Asn Trp Val Thr Ser Gly Asp Gly Ala Pro Tyr Pro Trp
Trp Met Arg Tyr Gln Pro Val Ser Tyr Ser Leu Asp Arg Ser Arg Ser Gly Thr Arg Ala Glu Phe Gln
Asp Met Val Asn Arg Cys Asn Ala Val Gly Val Gly Ile Tyr Val Asp Ala Val Ile Asn His Met Ser
Gly Gly Thr Gly Gly Thr Ser Ser Ala Gly Arg Ser Trp Ser Tyr His Asn Tyr Pro Gly Leu Tyr Gly
Pro Asn Asp Phe His Gln Pro Val Cys Ser Ile Thr Asn Tyr Gly Asp Ala Asn Asn Val Gln Arg Cys
Glu Leu Ser Gly Leu Gln Asp Leu Asp Thr Gly Ser Ala Tyr Val Arg Gly Lys Ile Ala Asp Tyr Leu
Val Asp Leu Val Asn Met Gly Val Lys Gly Phe Arg Val Asp Ala Ala Lys His Ile Ser Pro Thr Asp
Leu Gly Ala Ile Ile Asp Ala Val Asn Ser Arg Thr Gly Ala Asn Arg Pro Phe Trp Phe Leu Glu Val
Ile Gly Ala Ala Gly Glu Ala Val Gln Pro Asn Gln Tyr Phe Ser Leu Gly Gly Gly Gln Val Thr Val
Thr Glu Phe Asn Tyr Gly Lys Gln Ile Phe Gly Lys Phe Ala Gly Gly Gly Arg Leu Ala Glu Leu Arg
Ser Phe Gly Glu Thr Trp Gly Leu Met Pro Ser Ser Lys Ala Ile Ala Phe Ile Asp Asn His Asp Lys
Gln Arg Gly His Gly Gly Gly Gly Asn Tyr Leu Thr Tyr His His Gly Ser Thr Tyr Asp Leu Ala Asn
Ile Phe Met Leu Ala Trp Pro Tyr Gly Tyr Pro Ala Leu Met Ser Ser Tyr Ala Phe Asn Arg Ser Thr
Ala Tyr Asp Thr Ser Phe Gly Pro Pro His Asp Ser Gly Gly Ala Thr Arg Gly Pro Trp Asp Gly Gly
Gly Ser Gln Pro Ala Cys Phe Asn Gln Ser Ile Gly Gly Trp Val Cys Glu His Arg Trp Arg Gly Ile
Ala Asn Met Val Ala Phe Arg Asn Ala Thr Leu Pro Asn Trp Thr Val Thr Asp Trp Trp Asp Asn
Gly Asn Asn Gln Ile Ala Phe Gly Arg Gly Asp Lys Gly Phe Val Val Ile Asn Arg Glu Asp Ala Ala
Leu Thr Arg Asn Phe Lys Thr Ser Leu Pro Ala Gly Gln Tyr Cys Asp Val Ile Ser Gly Asp Phe Asn
Asn Gly Gln Cys Thr Gly His Val Val Thr Val Asp Ala Gly Gly Tyr Val Thr Leu Thr Ala Gly Pro
Asn Gly Ala Ala Ala Ile His Val Gly Ala Arg Leu Asp Gly Ala Ser Gln Pro Pro Thr Thr Ala Ser
Val Thr Phe Asn Ala Ser Ala Asp Thr Phe Trp Gly Gln Asn Leu Phe Val Val Gly Asn His Ser Ala
Leu Gly Asn Trp Ser Pro Ala Ala Ala Arg Pro Met Thr Trp Ile Ser Gly Ser Gly Thr Arg Gly Asn
Trp Arg Ala Val Leu Asn Leu Pro Ala Asn Thr Thr Tyr Gln Tyr Lys Phe Ile Lys Lys Asp Gly Ala
Gly Asn Val Val Trp Glu Gly Gly Gly Asn Arg Val Val Thr Thr Pro Ser Gly Gly Gly Ser Val Ser
Thr Gly Gly Asn Trp Gln

SEQ ID NO: 133
atgaataatgtgaaaaaagtatggttgtattattctataattgctaccttagttatttccttttttcacaccttttttcaacagcacaagctaatactgcacctg
tcaacggaacaatgatgcaatatttcgaatgggatttaccgaatgatgggacgctttggacgaaagtaaaaaatgaagctaccaatctttttcttcgct

FIGURE 16BBB

aggtattacagcgttatggctccctccagcatataaaggaacgagccaaagcgatgtcggatatggcgtgtacgatttatatgaccttggggaatt
taatcaaaaagggacgatccgaacgaaatacggaacaaaagcacaatatattcaagccatccaagctgccaaagccgcagggatgcaagtat
atgcagatgttgtatttaatcataaggcgggggctgacggcacagaatttgtcgatgcagttgaggtaaacccttctaatcgaaatcaagaaacat
ctggcacatatcaaattcaagcatggacaaaatttgattttcctggtcgtggaaacacatactccagcttcaaatggcgctggtatcattttgacggt
accgattgggatgaaagtcgtaaattaaatcgtatttacaaattccgcggtacaggaaaagcgtgggactgggaagtcgatacagaaaacgga
aactatgattatttaatgttcgctgatttagatatggatcaccctgaagttgtgacagagttaaaaaactggggaaaatggtatgtaaatacgacaaa
tgtagacggatttcgtttggatgccgtaaaacatattaaatacagcttttccctgactggctaacatatgtacgtaatcaaacaggaaaaaatttattt
gctgttggggaattttggagctatgacgtcaataagctgcataactacattacaaaaacaaatggatcgatgtcgttatttgatgcacctttgcataa
caacttttatatcgcttccaaatcgagtggatattttgacatgcgttatttattgaataatacattaatgaaagatcaaccttcactcgctgtaacacttgt
cgataaccatgatacacaaccaggtcaatctttacaatcatgggtagaagcttggtttaaaccgcttgcttacgcctttatttttaacaagacaagag
gggtatccttgcgtattttacggtgactattacggaatcccgaaatacaatattccgggattaaaaagtaaaattgatccgcttttaattgctcgtcgt
gattatgcttatggaacacaacgtgattacattgatcatcaagacattatcggatggacacgagaaggcattgatgcaaaaccgaactctggactt
gcggctttaattaccgacggccctggcggaagtaaatggatgtatgtcggtaaaaaacatgctgggaaagtgttttatgatttaactggaaatcga
agtgacacagtaacgattaatgcggacggttggggagaatttaaagtaaacggcggctccgtttcgatttgggtggctaaaacatcaaacgtca
catttacagtcaataacgccacaacaacaagtggacaaaacgtatatgttgttggcaacattccagagctaggcaattctttg

SEQ ID NO: 134

Met Asn Asn Val Lys Lys Val Trp Leu Tyr Tyr Ser Ile Ile Ala Thr Leu Val Ile Ser Phe Phe Thr Pro Phe Ser Thr Ala Gln Ala Asn Thr Ala Pro Val Asn Gly Thr Met Met Gln Tyr Phe Glu Trp Asp Leu Pro Asn Asp Gly Thr Leu Trp Thr Lys Val Lys Asn Glu Ala Thr Asn Leu Ser Ser Leu Gly Ile Thr Ala Leu Trp Leu Pro Pro Ala Tyr Lys Gly Thr Ser Gln Ser Asp Val Gly Tyr Gly Val Tyr Asp Leu Tyr Asp Leu Gly Glu Phe Asn Gln Lys Gly Thr Ile Arg Thr Lys Tyr Gly Thr Lys Ala Gln Tyr Ile Gln Ala Ile Gln Ala Ala Lys Ala Ala Gly Met Gln Val Tyr Ala Asp Val Val Phe Asn His Lys Ala Gly Ala Asp Gly Thr Glu Phe Val Asp Ala Val Glu Val Asn Pro Ser Asn Arg Asn Gln Glu Thr Ser Gly Thr Tyr Gln Ile Gln Ala Trp Thr Lys Phe Asp Phe Pro Gly Arg Gly Asn Thr Tyr Ser Ser Phe Lys Trp Arg Trp Tyr His Phe Asp Gly Thr Asp Trp Asp Glu Ser Arg Lys Leu Asn Arg Ile Tyr Lys Phe Arg Gly Thr Gly Lys Ala Trp Asp Trp Glu Val Asp Thr Glu Asn Gly Asn Tyr Asp Tyr Leu Met Phe Ala Asp Leu Asp Met Asp His Pro Glu Val Val Thr Glu Leu Lys Asn Trp Gly Lys Trp Tyr Val Asn Thr Thr Asn Val Asp Gly Phe Arg Leu Asp Ala Val Lys His Ile Lys Tyr Ser Phe Phe Pro Asp Trp Leu Thr Tyr Val Arg Asn Gln Thr Gly Lys Asn Leu Phe Ala Val Gly Glu Phe Trp Ser Tyr Asp Val Asn Lys Leu His Asn Tyr Ile Thr Lys Thr Asn Gly Ser Met Ser Leu Phe Asp Ala Pro Leu His Asn Asn Phe Tyr Ile Ala Ser Lys Ser Ser Gly Tyr Phe Asp Met Arg Tyr Leu Leu Asn Asn Thr Leu Met Lys Asp Gln Pro Ser Leu Ala Val Thr Leu Val Asp Asn His Asp Thr Gln Pro Gly Gln Ser Leu Gln Ser Trp Val Glu Ala Trp Phe Lys Pro Leu Ala Tyr Ala Phe Ile Leu Thr Arg Gln Glu Gly Tyr Pro Cys Val Phe Tyr Gly Asp Tyr Tyr Gly Ile Pro Lys Tyr Asn Ile Pro Gly Leu Lys Ser Lys Ile Asp Pro Leu Leu Ile Ala Arg Arg Asp Tyr Ala Tyr Gly Thr Gln Arg Asp Tyr Ile Asp His Gln Asp Ile Ile Gly Trp Thr Arg Glu Gly Ile Asp Ala Lys Pro Asn Ser Gly Leu Ala Ala Leu Ile Thr Asp Gly Pro Gly Gly Ser Lys Trp Met Tyr Val Gly Lys Lys His Ala Gly Lys Val Phe Tyr Asp Leu Thr Gly Asn Arg Ser Asp Thr Val Thr Ile Asn Ala Asp Gly Trp Gly Glu Phe Lys Val Asn Gly Gly Ser Val Ser Ile Trp Val Ala Lys Thr Ser Asn Val Thr Phe Thr Val Asn Asn Ala Thr Thr Thr Ser Gly Gln Asn Val Tyr Val Val Gly Asn Ile Pro Glu Leu Gly Asn Ser Leu

SEQ ID NO: 135

gtgacaggcaccccgtctttatacattcctccacataaaataaccatacagctttcaaatttgttgaaatgtataaaaataaaaaatagtattgtaagc
gttaacatccgtcattataataacttcaaacgcgtttatgttttaatgcaaacgtttgcatcctcatttttatttaaagaaaggatgtgtgtgcatgaattatt
tgaaaaaagtgtggttgtattacgctatcgtcgctaccttaatcatttccttttcttacgccctttttcaactgcacaagccaacactgcaccagtcaacg
gaacgatgatgcaatatttcgaatgggatttaccgaatgatggcacactttggacgaaagtaaaaaacgaagcaagcagccttttcttctttaggtat
tactgcgttatggttaccacctgcatacaaaggaacgagccaaggggatgtcgggtatggcgtgtacgatttgtatgacttaggagaatttaatca
aaaagggacgattcgaacgaaatacggaacaaaaacgcaatatttacaagccattcaagcggcaaaaagcgctggcatgcaagtatacgctg
atgtcgtatttaatcacaaggcgggggcagatagtacagaatgggttgacgcagtcgaagtgaatccttctaatcgaaaccaagaaacatctgg

FIGURE 16CCC

cacatatcaaattcaagcatggacaaaatttgatttccctgaccgtgggaacacatactcaagctttaaatggcgctggtatcattttgacggtacg
gattgggatgaaagtcgaaaactaaatcgcatttacaaatttcgtggcacaggaaaagcatgggattgggaagtagacacagagaacggaaac
tatgactacttaatgtttgctgatttagatatggatcaccctgaagtcgtgacagagctaaaaaactggggaacatggtacgtcaatacgacaaatg
tcgatgggtttcgcttagatgcagtaaagcatattaaatatagctttttcccagattggttaacatatgtgcgctcacaaacacaaaaaaatctgtttg
cagtaggagaattttggagctacgatgtcaataaactgcataactacattacaaaaacaagtggaaccatgtcgttatttgatgcgccacttcataa
caacttttacactgcttcaaaatctagcgggtattttgacatgcgctatttgttaaataatacgttgatgaaagaccagccttctcttgcggtcacactc
gttgataatcatgacacgcaaccgggacaatctttacaatcatgggtagagccttggtttaagccgcttgcttatgcctttatttttgacaagacaaga
aggatatccttgcgtattttacggcgactattacggcatccctaaatacaatattccgggattgaaaagtaaaatcgatccgcttctcattgcccgta
gagactacgcatacggaacacaacgtgattatattgaccatcaagacattattggatggacacgggaaggaattgactcaaaaccgaactctgg
acttgcggctttaattactgacggtcctggtggaagtaaatggatgtatgtaggtaaaaagcatgctggaaaagtgttttacgatctcactggaaat
cgaagcgatacggtaacgattaatgcagacggctggggagagtttaaagtaaacggtggctccgtttccatttgggttgccaaaacatcacaag
tcacgtttaccgtcaacaatgcgacaacgacaagcggacaaaatgtgtatgtcgttggcaacattccagagctcggaaattggaacacagcaaa
cgcaatcaaaatgaccccatcttcttatccaacgtggaaaacaaccattgctcttccacaaggaaaagcaattggcggcgtacgccatggcccttt
ga

SEQ ID NO: 136

Val Thr Gly Thr Pro Ser Leu Tyr Ile Pro Pro His Lys Ile Thr Ile Gln Leu Ser Asn Leu Leu Lys Cys
Ile Lys Ile Lys Asn Ser Ile Val Ser Val Asn Ile Arg His Tyr Asn Asn Phe Lys Arg Val Tyr Val Leu
Met Gln Thr Phe Ala Ser Ser Phe Tyr Leu Lys Lys Gly Cys Val Cys Met Asn Tyr Leu Lys Lys Val
Trp Leu Tyr Tyr Ala Ile Val Ala Thr Leu Ile Ile Ser Phe Leu Thr Pro Phe Ser Thr Ala Gln Ala Asn
Thr Ala Pro Val Asn Gly Thr Met Met Gln Tyr Phe Glu Trp Asp Leu Pro Asn Asp Gly Thr Leu
Trp Thr Lys Val Lys Asn Glu Ala Ser Ser Leu Ser Ser Leu Gly Ile Thr Ala Leu Trp Leu Pro Pro
Ala Tyr Lys Gly Thr Ser Gln Gly Asp Val Gly Tyr Gly Val Tyr Asp Leu Tyr Asp Leu Gly Glu Phe
Asn Gln Lys Gly Thr Ile Arg Thr Lys Tyr Gly Thr Lys Thr Gln Tyr Leu Gln Ala Ile Gln Ala Ala
Lys Ser Ala Gly Met Gln Val Tyr Ala Asp Val Val Phe Asn His Lys Ala Gly Ala Asp Ser Thr Glu
Trp Val Asp Ala Val Glu Val Asn Pro Ser Asn Arg Asn Gln Glu Thr Ser Gly Thr Tyr Gln Ile Gln
Ala Trp Thr Lys Phe Asp Phe Pro Asp Arg Gly Asn Thr Tyr Ser Ser Phe Lys Trp Arg Trp Tyr His
Phe Asp Gly Thr Asp Trp Asp Glu Ser Arg Lys Leu Asn Arg Ile Tyr Lys Phe Arg Gly Thr Gly Lys
Ala Trp Asp Trp Glu Val Asp Thr Glu Asn Gly Asn Tyr Asp Tyr Leu Met Phe Ala Asp Leu Asp
Met Asp His Pro Glu Val Val Thr Glu Leu Lys Asn Trp Gly Thr Trp Tyr Val Asn Thr Thr Asn Val
Asp Gly Phe Arg Leu Asp Ala Val Lys His Ile Lys Tyr Ser Phe Phe Pro Asp Trp Leu Thr Tyr Val
Arg Ser Gln Thr Gln Lys Asn Leu Phe Ala Val Gly Glu Phe Trp Ser Tyr Asp Val Asn Lys Leu His
Asn Tyr Ile Thr Lys Thr Ser Gly Thr Met Ser Leu Phe Asp Ala Pro Leu His Asn Asn Phe Tyr Thr
Ala Ser Lys Ser Ser Gly Tyr Phe Asp Met Arg Tyr Leu Leu Asn Asn Thr Leu Met Lys Asp Gln
Pro Ser Leu Ala Val Thr Leu Val Asp Asn His Asp Thr Gln Pro Gly Gln Ser Leu Gln Ser Trp Val
Glu Pro Trp Phe Lys Pro Leu Ala Tyr Ala Phe Ile Leu Thr Arg Gln Glu Gly Tyr Pro Cys Val Phe
Tyr Gly Asp Tyr Tyr Gly Ile Pro Lys Tyr Asn Ile Pro Gly Leu Lys Ser Lys Ile Asp Pro Leu Leu Ile
Ala Arg Arg Asp Tyr Ala Tyr Gly Thr Gln Arg Asp Tyr Ile Asp His Gln Asp Ile Ile Gly Trp Thr
Arg Glu Gly Ile Asp Ser Lys Pro Asn Ser Gly Leu Ala Ala Leu Ile Thr Asp Gly Pro Gly Gly Ser
Lys Trp Met Tyr Val Gly Lys Lys His Ala Gly Lys Val Phe Tyr Asp Leu Thr Gly Asn Arg Ser Asp
Thr Val Thr Ile Asn Ala Asp Gly Trp Gly Glu Phe Lys Val Asn Gly Gly Ser Val Ser Ile Trp Val
Ala Lys Thr Ser Gln Val Thr Phe Thr Val Asn Asn Ala Thr Thr Thr Ser Gly Gln Asn Val Tyr Val
Val Gly Asn Ile Pro Glu Leu Gly Asn Trp Asn Thr Ala Asn Ala Ile Lys Met Thr Pro Ser Ser Tyr
Pro Thr Trp Lys Thr Thr Ile Ala Leu Pro Gln Gly Lys Ala Ile Gly Gly Val Arg His Gly Pro

SEQ ID NO: 137

gtgggacgggcaggcttggcgcatcactcgaacacttccgccaaggggacatacgggtcacctctcgaactgcgtccggatcgccccgccgt
ggccggggcggtcgagcttgaagatgtccagcggggagccgccgccgaggatcaccccggcggcgtactcgcccagggcggggctcag
cttgaagccgtggccggagccgcctcccaggagccagacgttggaggcccgcggatggcggtcgaggaggaggtggccgtcggggctgt
tctcgtactggcagacgcgggtctcgaccagcggcgcgtccttcagggccgggaaccggcgggccacctcggccgggccgcttccagca

FIGURE 16DDD

gggccggggtgatcgtccgctcgcccgccgtgggatcgatgggctcgccccgggtgtcgtccgccaccttgaagccgcggtgctcgttgcc
ggggatgccgtagtagatccgctcgccgagatcgacccagaccggacagccgccctcctggaagcgcgggtcgcccggcggcgtgccga
agaagaacacctcctggcgggtgttgcggaggaaccgctcaccgatcacgtccgggaacagcccggccagccagggaccgcaggcgaag
acgtagaggtcggccgcgagagtggagccgtccgaaaggtgaagccgctccaagggccccgggaccatggcggcctgccggtactcccc
gccctcgccctggaacagctccaccacggtccggcaggcgcgccgggcgaacagggcgccggcttcctcctcgtaccagatcgtgcggac
gccgtcgaaatcgacctgggggaagcggctccgggcctcccccctgagacagctcggcgaccggcagcccccgcgtcctccagaaaaggaa
gggagtcgcggacgtagctgtcgtcctcgccgcacatccagaggaccccggtccttttgtacagccggtaaccggactggacttcggcgtccc
gccagagctcgaaggagcgggcgacccactccacgtacagacggtcgggtccgtaggcgccgcggatgatccgcgtctcgccaccggag
ctggagcgggagtgccccggaccccaggcgtccaggagggtcacccgggctccgcggcggaggagatgcagggcggtccagccgccg
aaggcgccggcgccgacgacggcgatatggggatgggagggcatggcgggcgtaaggttatcgcagcccgatccttcgctggcatcccat
ctccgaccggagtatcctggaaaattcgaagaaggagatcgacatgcaatcgaacggaaacgtga

SEQ ID NO: 138
Val Gly Arg Ala Gly Leu Ala His His Ser Asn Thr Ser Ala Lys Gly Thr Tyr Gly Ser Pro Leu Glu Leu Arg Pro Asp Arg Pro Ala Val Ala Gly Ala Val Glu Leu Glu Asp Val Gln Arg Gly Ala Ala Ala Glu Asp His Pro Gly Gly Val Leu Ala Gln Gly Gly Ala Gln Leu Glu Ala Val Ala Gly Ala Ala Ser Gln Glu Pro Asp Val Gly Gly Pro Arg Met Ala Val Glu Glu Glu Val Ala Val Gly Ala Val Leu Val Leu Ala Asp Ala Gly Leu Asp Gln Arg Arg Val Leu Gln Gly Arg Glu Pro Ala Gly His Leu Gly Pro Gly Arg Phe Gln Gln Gly Arg Gly Asp Arg Pro Leu Ala Arg Arg Gly Ile Asp Gly Leu Ala Pro Gly Val Val Arg His Leu Glu Ala Ala Val Leu Val Ala Gly Asp Ala Val Val Asp Pro Leu Ala Glu Ile Asp Pro Asp Arg Thr Ala Ala Leu Leu Glu Ala Arg Val Ala Arg Arg Arg Ala Glu Glu Glu His Leu Leu Ala Gly Val Ala Glu Glu Pro Leu Thr Asp His Val Arg Glu Gln Pro Gly Gln Pro Gly Thr Ala Gly Glu Asp Val Glu Val Gly Arg Glu Ser Gly Ala Val Arg Lys Val Lys Pro Leu Gln Gly Pro Arg Asp His Gly Gly Leu Pro Val Leu Pro Ala Leu Ala Leu Glu Gln Leu His His Gly Pro Ala Gly Ala Pro Gly Glu Gln Gly Ala Gly Phe Leu Leu Val Pro Asp Arg Ala Asp Ala Val Glu Ile Asp Leu Gly Glu Ala Ala Pro Gly Leu Pro Leu Arg Gln Leu Gly Asp Arg Gln Pro Arg Val Leu Gln Lys Arg Lys Gly Val Ala Asp Val Ala Val Val Leu Ala Ala His Pro Glu Asp Pro Gly Pro Phe Val Gln Pro Val Thr Gly Leu Asp Phe Gly Val Pro Pro Glu Leu Glu Gly Ala Gly Asp Pro Leu His Val Gln Thr Val Gly Ser Val Gly Ala Ala Asp Asp Pro Arg Leu Ala Thr Gly Ala Gly Ala Gly Val Pro Arg Thr Pro Gly Val Gln Glu Gly His Pro Gly Ser Ala Ala Glu Glu Met Gln Gly Gly Pro Ala Ala Glu Gly Ala Gly Ala Asp Asp Gly Asp Met Gly Met Gly Gly His Gly Gly Arg Lys Val Ile Ala Ala Arg Ser Phe Ala Gly Ile Pro Ser Pro Thr Gly Val Ser Trp Lys Ile Arg Arg Arg Arg Ser Thr Cys Asn Arg Thr Glu Thr

SEQ ID NO: 139
atgaaaacattcaaccttaaacccacactttaccttaactttgctgctgagttcgccggtattggcggcacaaaatggaaaetatgatgcagtatttc
cattggtatgtgccaaatgacggcgcaatctggaacacaagttgaaaactatgcgccagcactatccgaaaacggtttacagcgctgtggttgc
caccagcatataaaggagcaggtggtagcaacgacgttggttacggtgtttacgatatgtatgacttaggggagtttgatcaaaaaggatcggta
cgaactaagtacggcaccaaagaccaatatctaaatgccatcaaagcagcacacaaaaacaatatccaaatttatggtgacgtagtgttcaacca
tcgtggcggtgcagatggcaagtcgtgggtcgataccaagcgtgtggattggaataaccgcaatattgaacttggcgataaatggattgaagca
tgggttgaatttagcttcccaggacgtaacgataaatactcagacttccattggacgtggtatcactttgatggcgtcgattgggatgacgcaggta
aagagaaagcgatctttaaattcaaaggtgatggtaaagcatgggattgggaagtcagttctgaaaaaggcaactatgactacctcatgtacgca
gacttagacatggatcacccagaagtgaagcaagagctgaaagattggggtgaatggtacttaaacatgacgggtgttgatggcttccgaatgg
atgcagtgaaacacatcaaatatcagtacctacaagagtggatcgattacttgcgtaagaaaacgggcaaagagctctttaccgttggtgagtac
tggaactacgacgtgaacaatctgcacaacttttatgactaagacttctggcagcatgtcattgtttgatgcgccttacatatgaacttctataacgct
tcacgctctggtggcaacttgatatgcgccgaatcatggatggcaccttgatgaaagacaacccagtgaaagcagtaacactggttgagaacc
atgatacgcaaccactacaggccttagagtctccggtggattggtggttcaaaccacttgcgtacgcgttcattttgcttcgtgaggaaggttatcc
gtcagtcttctacgcagattactacggtgcgcaatacagcgataaagggcacgatatcaacatggtgaaagtgccttacattgagcaattggtga
aagcgcgtaaagattatgcttatggtaaacaacattcttaccttgaccactgggatgtgattggttggacacgagaaggggatgcggaacatccg

2013201807 18 Mar 2013

FIGURE 16EEE

aactctatggcggttatcatgagtgatggtcctggcggaacaaagtggatgtacacaggttcaccgagcacacgttatgtcgataaactaggtatt
cgtaccgaagaagtatggactaacgctagtggatgggccgaattcccagtgaacggcggatcggtttctgtttgggttggcgttaaataa

SEQ ID NO: 140
Met Lys Thr Phe Asn Leu Lys Pro Thr Leu Leu Pro Leu Thr Leu Leu Leu Ser Ser Pro Val Leu Ala
Ala Gln Asn Gly Thr Met Met Gln Tyr Phe His Trp Tyr Val Pro Asn Asp Gly Ala Leu Trp Thr Gln
Val Glu Asn Asn Ala Pro Ala Leu Ser Asp Asn Gly Phe Thr Ala Leu Trp Leu Pro Pro Ala Tyr Lys
Gly Ala Gly Gly Ser Asn Asp Val Gly Tyr Gly Val Tyr Asp Met Tyr Asp Leu Gly Glu Phe Asp
Gln Lys Gly Ser Val Arg Thr Lys Tyr Gly Thr Lys Asp Gln Tyr Leu Asn Ala Ile Lys Ala Ala His
Lys Asn Asn Ile Gln Ile Tyr Gly Asp Val Val Phe Asn His Arg Gly Gly Ala Asp Gly Lys Ser Trp
Val Asp Thr Lys Arg Val Asp Trp Asn Asn Arg Asn Ile Glu Leu Gly Asp Lys Trp Ile Glu Ala Trp
Val Glu Phe Ser Phe Pro Gly Arg Asn Asp Lys Tyr Ser Asp Phe His Trp Thr Trp Tyr His Phe Asp
Gly Val Asp Trp Asp Asp Ala Gly Lys Glu Lys Ala Ile Phe Lys Phe Lys Gly Asp Gly Lys Ala Trp
Asp Trp Glu Val Ser Ser Glu Lys Gly Asn Tyr Asp Tyr Leu Met Tyr Ala Asp Leu Asp Met Asp
His Pro Glu Val Lys Gln Glu Leu Lys Asp Trp Gly Glu Trp Tyr Leu Asn Met Thr Gly Val Asp Gly
Phe Arg Met Asp Ala Val Lys His Ile Lys Tyr Gln Tyr Leu Gln Glu Trp Ile Asp Tyr Leu Arg Lys
Lys Thr Gly Lys Glu Leu Phe Thr Val Gly Glu Tyr Trp Asn Tyr Asp Val Asn Asn Leu His Asn
Phe Met Thr Lys Thr Ser Gly Ser Met Ser Leu Phe Asp Ala Pro Leu His Met Asn Phe Tyr Asn Ala
Ser Arg Ser Gly Gly Asn Phe Asp Met Arg Arg Ile Met Asp Gly Thr Leu Met Lys Asp Asn Pro
Val Lys Ala Val Thr Leu Val Glu Asn His Asp Thr Gln Pro Leu Gln Ala Leu Glu Ser Pro Val Asp
Trp Trp Phe Lys Pro Leu Ala Tyr Ala Phe Ile Leu Leu Arg Glu Glu Gly Tyr Pro Ser Val Phe Tyr
Ala Asp Tyr Tyr Gly Ala Gln Tyr Ser Asp Lys Gly His Asp Ile Asn Met Val Lys Val Pro Tyr Ile
Glu Gln Leu Val Lys Ala Arg Lys Asp Tyr Ala Tyr Gly Lys Gln His Ser Tyr Leu Asp His Trp Asp
Val Ile Gly Trp Thr Arg Glu Gly Asp Ala Glu His Pro Asn Ser Met Ala Val Ile Met Ser Asp Gly
Pro Gly Gly Thr Lys Trp Met Tyr Thr Gly Ser Pro Ser Thr Arg Tyr Val Asp Lys Leu Gly Ile Arg
Thr Glu Glu Val Trp Thr Asn Ala Ser Gly Trp Ala Glu Phe Pro Val Asn Gly Gly Ser Val Ser Val
Trp Val Gly Val Lys

SEQ ID NO: 141
atgaaaccaataaataccctactcatatccgccccttgctgtttgttctttcagttccgcgacttacgccgatactattttgcacgcgttcaattggaagt
attcagatgtgacggccaacgcgaatcaaattgctcaagctggttataagaaagtgcttgttgcgcctgcaatgaaatcgagtggcagccaatgg
tgggctcgctatcaacctcaagatctacgcactatcgattctcctttgggcaataaacaagatttagccgcaatgattgccgcactcaaaggtgtg
ggcgtcgatgtgtatgccgatgtggtactcaaccatatggcgaatgaaagctggaagcgaagtgacttgaattaccctggcacagaagtgctaa
acgattatgctagccgttcaagctactatgctgaccagactctgtttggcaacctagcacaaggttatgtgtcagcgaacgactttcatccagcgg
gctgtatttcagattggaacgaccctggtcatgttcagtattggcgtttgtgtggcgcagatggtgatgtaggtttacctgaccttgatccaaacaac
tgggtggtttcacaacagcgtttgtatctgaaagcgctaaaagatatgggcatcaaagggttccgaattgatgcagtgaagcacatgagccaata
ccaaatcgatcaggtattcacgtctgaaattactgcgaacatgcatgtgtttggtgaagtgattactagcggtgcagcagcggtatagcgcgtatg
aatcgttcttagcgccttacctgaatatatctaatcactctgcctacgatttcccgctgtttgcatcgattcgctcggcatttctatgggggcggttt
aaatcaactgcatgatcctaaagcgtacggtcaggcacttgatgataatcgctcgatcacctttgcgatcacacatgatattccaaccaatgacgg
cttccgctaccaaattatggacccacaagacgagcagcttgcttacgcgtatatccttggtaaagacggtggcacgccgctgatctacagtgatg
atcttcctgattctgaagacaaggataacggtcgttggggcaatgtttggaacagttcgacaatgaaaaacatgttgagcttccataacgcgatgc
aaggcaaaacaatgacgatgatttctagcgaccattgcactttgttgtttaagcgtggcaaagaaggtgttgtgggtattaacaagtgtggtgaaa
cgcgtggcgtgacggttgatacctaccaacatgagtttaattggcatgttcaatacaaagacgtgttaagcagcgcaacagaaaccgtgacttct
cgttaccatacgttcaatctaccaccacgcagtgcgcgtatgtttaagctgtag

SEQ ID NO: 142
Met Lys Pro Ile Asn Thr Leu Leu Ile Ser Ala Leu Ala Val Cys Ser Phe Ser Ser Ala Thr Tyr Ala
Asp Thr Ile Leu His Ala Phe Asn Trp Lys Tyr Ser Asp Val Thr Ala Asn Ala Asn Gln Ile Ala Gln
Ala Gly Tyr Lys Lys Val Leu Val Ala Pro Ala Met Lys Ser Ser Gly Ser Gln Trp Trp Ala Arg Tyr
Gln Pro Gln Asp Leu Arg Thr Ile Asp Ser Pro Leu Gly Asn Lys Gln Asp Leu Ala Ala Met Ile Ala

FIGURE 16FFF

Ala Leu Lys Gly Val Gly Val Asp Val Tyr Ala Asp Val Val Leu Asn His Met Ala Asn Glu Ser Trp Lys Arg Ser Asp Leu Asn Tyr Pro Gly Thr Glu Val Leu Asn Asp Tyr Ala Ser Arg Ser Ser Tyr Tyr Ala Asp Gln Thr Leu Phe Gly Asn Leu Ala Gln Gly Tyr Val Ser Ala Asn Asp Phe His Pro Ala Gly Cys Ile Ser Asp Trp Asn Asp Pro Gly His Val Gln Tyr Trp Arg Leu Cys Gly Ala Asp Gly Asp Val Gly Leu Pro Asp Leu Asp Pro Asn Asn Trp Val Val Ser Gln Gln Arg Leu Tyr Leu Lys Ala Leu Lys Asp Met Gly Ile Lys Gly Phe Arg Ile Asp Ala Val Lys His Met Ser Gln Tyr Gln Ile Asp Gln Val Phe Thr Ser Glu Ile Thr Ala Asn Met His Val Phe Gly Glu Val Ile Thr Ser Gly Gly Ala Gly Asn Ser Gly Tyr Glu Ser Phe Leu Ala Pro Tyr Leu Asn Asn Thr Asn His Ser Ala Tyr Asp Phe Pro Leu Phe Ala Ser Ile Arg Ser Ala Phe Ser Met Gly Gly Gly Leu Asn Gln Leu His Asp Pro Lys Ala Tyr Gly Gln Ala Leu Asp Asp Asn Arg Ser Ile Thr Phe Ala Ile Thr His Asp Ile Pro Thr Asn Asp Gly Phe Arg Tyr Gln Ile Met Asp Pro Gln Asp Glu Gln Leu Ala Tyr Ala Tyr Ile Leu Gly Lys Asp Gly Gly Thr Pro Leu Ile Tyr Ser Asp Asp Leu Pro Asp Ser Glu Asp Lys Asp Asn Gly Arg Trp Gly Asn Val Trp Asn Ser Ser Thr Met Lys Asn Met Leu Ser Phe His Asn Ala Met Gln Gly Lys Thr Met Thr Met Ile Ser Ser Asp His Cys Thr Leu Leu Phe Lys Arg Gly Lys Glu Gly Val Val Gly Ile Asn Lys Cys Gly Glu Thr Arg Gly Val Thr Val Asp Thr Tyr Gln His Glu Phe Asn Trp His Val Gln Tyr Lys Asp Val Leu Ser Ser Ala Thr Glu Thr Val Thr Ser Arg Tyr His Thr Phe Asn Leu Pro Pro Arg Ser Ala Arg Met Phe Lys Leu

SEQ ID NO: 143

atgccaaagagcacttttaccaaatccataacaaaatcacttcttgctacttccgttgttgtaagcttattgcctgcctacgcacaggccgacactat
cttgcatgccttttaactggaaatacagcgacattacccgccaagcagagcaaattgcgcaagctggttataaaaaagtactgatttcaccgccgc
tgaagtccacaggcccacaatggtgggcacgttaccaaccacaggacattcgagtgattgactcccctgtcggcaacaagcaagatttacaag
ccctcattgcagccttaaaggcacaaggcgttgaagtatacgcagacatcgtactcaaccacatggccaacgaaagctggaaacgagacgatc
tgaactacccgggaagtgatttacttacccaatacagccaaaatatggcttacatgaaccagcaaaaattgtttggagatttagagcaaaatcagtt
ctctgccaatgattttcacccggctggctgcattactgattggagtaacccggggcatgttcaatactggcgcttatgtggtggtaatggtgacact
gggttacctgatcttgatcctaactcgtgggtgatcgatcaacaaaaacgttatttacgtgctttgaaagacatgggaataaagggcttccgagttg
atgcggtaaaacacatgagcgattaccaaatcaaccaagtgtttacgccagacatcatcgcaggcttacatgtatttggtgaagtgatcaccagtg
gtggcaagggcagcaatgactaccactcttttctggaaccgtatttaaataacaccaatcacgccgcgtatgacttcccgctatttgcctctatccg
aaatgcatttagttatcatggcagcttgtctcaattacatgatccacaagcttacgggcaagcacttcctaacgacagagccattactttcaccatca
ctcacgacattccaaccaatgatggtttccgttaccaaatcatggatccaaccagtgaaaaactcgcgtacgcgtacattctaggcaaagatggg
ggtagcccacttatctatagcgatgctttagacccaagtgaagataagataagggccgctggcgtgatgtatggaaccaagaatacatggttaa
catgatcagcttccacaacaaggtgcaaggtaaaagcatggaggtcatgtacagcgatcaatgcttgctggtctttaaacgtgaaaaacaaggct
tagtcggtattaataagtgcgctgaaagccgtacctacaccatagatacccatcgttttgaatttaactggtaccaaccgtacaacgacacattaag
ccagcacagcgagacctttagcagccgttatcatgctctgaccattccggcgcaaacagcacgaatgttggcgctataa

SEQ ID NO: 144

Met Pro Lys Ser Thr Phe Thr Lys Ser Ile Thr Lys Ser Leu Leu Ala Thr Ser Val Val Val Ser Leu Leu Pro Ala Tyr Ala Gln Ala Asp Thr Ile Leu His Ala Phe Asn Trp Lys Tyr Ser Asp Ile Thr Arg Gln Ala Glu Gln Ile Ala Gln Ala Gly Tyr Lys Lys Val Leu Ile Ser Pro Pro Leu Lys Ser Thr Gly Pro Gln Trp Trp Ala Arg Tyr Gln Pro Gln Asp Ile Arg Val Ile Asp Ser Pro Val Gly Asn Lys Gln Asp Leu Gln Ala Leu Ile Ala Ala Leu Lys Ala Gln Gly Val Glu Val Tyr Ala Asp Ile Val Leu Asn His Met Ala Asn Glu Ser Trp Lys Arg Asp Asp Leu Asn Tyr Pro Gly Ser Asp Leu Leu Thr Gln Tyr Ser Gln Asn Met Ala Tyr Met Asn Gln Gln Lys Leu Phe Gly Asp Leu Glu Gln Asn Gln Phe Ser Ala Asn Asp Phe His Pro Ala Gly Cys Ile Thr Asp Trp Ser Asn Pro Gly His Val Gln Tyr Trp Arg Leu Cys Gly Gly Asn Gly Asp Thr Gly Leu Pro Asp Leu Asp Pro Asn Ser Trp Val Ile Asp Gln Gln Lys Arg Tyr Leu Arg Ala Leu Lys Asp Met Gly Ile Lys Gly Phe Arg Val Asp Ala Val Lys His Met Ser Asp Tyr Gln Ile Asn Gln Val Phe Thr Pro Asp Ile Ile Ala Gly Leu His Val Phe Gly Glu Val Ile Thr Ser Gly Gly Lys Gly Ser Asn Asp Tyr His Ser Phe Leu Glu Pro Tyr Leu Asn Asn Thr Asn His Ala Ala Tyr Asp Phe Pro Leu Phe Ala Ser Ile Arg Asn Ala Phe Ser Tyr His Gly Ser Leu Ser Gln Leu His Asp Pro Gln Ala Tyr Gly Gln Ala Leu Pro Asn Asp Arg Ala Ile Thr Phe Thr Ile Thr His Asp

FIGURE 16GGG

Ile Pro Thr Asn Asp Gly Phe Arg Tyr Gln Ile Met Asp Pro Thr Ser Glu Lys Leu Ala Tyr Ala Tyr Ile Leu Gly Lys Asp Gly Gly Ser Pro Leu Ile Tyr Ser Asp Ala Leu Asp Pro Ser Glu Asp Lys Asp Lys Gly Arg Trp Arg Asp Val Trp Asn Gln Glu Tyr Met Val Asn Met Ile Ser Phe His Asn Lys Val Gln Gly Lys Ser Met Glu Val Met Tyr Ser Asp Gln Cys Leu Leu Val Phe Lys Arg Glu Lys Gln Gly Leu Val Gly Ile Asn Lys Cys Ala Glu Ser Arg Thr Tyr Thr Ile Asp Thr His Arg Phe Glu Phe Asn Trp Tyr Gln Pro Tyr Asn Asp Thr Leu Ser Gln His Ser Glu Thr Phe Ser Ser Arg Tyr His Ala Leu Thr Ile Pro Ala Gln Thr Ala Arg Met Leu Ala Leu

SEQ ID NO: 145

atgttgaaaaggattacggtagtctgtttattatttattttgctttttcctaatatatatgggaggaataaggcggaagcagcaacgataaataatgga
acattaatgcagtattttgagtggtacgctccgaatgatgggaatcattggaatcgtttgcgttatgatgctgaaagtttagctcataagggaatcac
atctgtatggataccacctgcatataaagggacttcgcaaaatgatgtagggtatggggcctatgatttatacgatttaggggagttcaatcaaaaa
ggaacggtgcggacgaaatatgggacaaaggcacagttgaaatctgcaattgacgctttacataagcaaaacatcgacgtatacggtgatgta
gttatgaatcataaaggtggggctgattatactgaaaccgtaacagctgttgaggtagaccgtaacaatcgaaatattgaagtatcaggtgattatg
aaattagtgcgtggacgggttttaactttccagggcgcagagatgcttattctaatttcaaatggaaatggtatcattttgacggaacggattgggat
gaaggaaggaaattaaaccgaatttataaatttaggggtataggtaaagcgtgggactgggaagtgtctagcgaaaatggaaattatgattatttg
atgtatgcagatcttgattttgatcatccagatgttgcgaatgaaatgaaaagttggggaacgtggtatgcgaatgaattaaatttagatggatttcgt
ttagatgctgttaaacatattgatcatgaatatttacgcgattgggtaaatcatgtcagacagcaaacggggaaagaaatgtttacggtggctgaat
attggcaaaatgatatccagactttaaacaattatttggcgaaagtcaattataatcaatctgtatttgatgcaccgcttcattacaattttcattatgctt
caacaggaaatgggaattatgatatgagaaatattttaaatggaacagtaatgaaaaatcatcctgcactcgcagttactctcgttgagaatcatga
ttctcaacctgggcaatcattggaatctgtagtaagtccgtggtttaagccgctggcatatgcatttattttaactcgtgcagagggctatccttcagt
ttttatggtgattactatgggacaagcggaaatagtagttatgaaattccagcgttaaaagataaaattgatccaattttgacggcacgaaaaaact
ttgcatatggtacgcagcgtgattatttagaccatccagatgtgattggctggacaagagaaggagatagtgtacatgctaagtctggtttagcgg
cattaatctccgatggaccaggaggatcaaagtggatggatgttggaaagaataacgctggggaagtatggtacgatattacgggtaatcaaac
aaatactgtaacaattaataagatggatcggggcaattccatgtaagtggaggctctgtttctatatatgttcaacagtaa

SEQ ID NO: 146

Met Leu Lys Arg Ile Thr Val Val Cys Leu Leu Phe Ile Leu Leu Phe Pro Asn Ile Tyr Gly Arg Asn Lys Ala Glu Ala Ala Thr Ile Asn Asn Gly Thr Leu Met Gln Tyr Phe Glu Trp Tyr Ala Pro Asn Asp Gly Asn His Trp Asn Arg Leu Arg Tyr Asp Ala Glu Ser Leu Ala His Lys Gly Ile Thr Ser Val Trp Ile Pro Pro Ala Tyr Lys Gly Thr Ser Gln Asn Asp Val Gly Tyr Gly Ala Tyr Asp Leu Tyr Asp Leu Gly Glu Phe Asn Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly Thr Lys Ala Gln Leu Lys Ser Ala Ile Asp Ala Leu His Lys Gln Asn Ile Asp Val Tyr Gly Asp Val Val Met Asn His Lys Gly Gly Ala Asp Tyr Thr Glu Thr Val Thr Ala Val Glu Val Asp Arg Asn Asn Arg Asn Ile Glu Val Ser Gly Asp Tyr Glu Ile Ser Ala Trp Thr Gly Phe Asn Phe Pro Gly Arg Arg Asp Ala Tyr Ser Asn Phe Lys Trp Lys Trp Tyr His Phe Asp Gly Thr Asp Trp Asp Glu Gly Arg Lys Leu Asn Arg Ile Tyr Lys Phe Arg Gly Ile Gly Lys Ala Trp Asp Trp Glu Val Ser Ser Glu Asn Gly Asn Tyr Asp Tyr Leu Met Tyr Ala Asp Leu Asp Phe Asp His Pro Asp Val Ala Asn Glu Met Lys Ser Trp Gly Thr Trp Tyr Ala Asn Glu Leu Asn Leu Asp Gly Phe Arg Leu Asp Ala Val Lys His Ile Asp His Glu Tyr Leu Arg Asp Trp Val Asn His Val Arg Gln Gln Thr Gly Lys Glu Met Phe Thr Val Ala Glu Tyr Trp Gln Asn Asp Ile Gln Thr Leu Asn Asn Tyr Leu Ala Lys Val Asn Tyr Asn Gln Ser Val Phe Asp Ala Pro Leu His Tyr Asn Phe His Tyr Ala Ser Thr Gly Asn Gly Asn Tyr Asp Met Arg Asn Ile Leu Asn Gly Thr Val Met Lys Asn His Pro Ala Leu Ala Val Thr Leu Val Glu Asn His Asp Ser Gln Pro Gly Gln Ser Leu Glu Ser Val Val Ser Pro Trp Phe Lys Pro Leu Ala Tyr Ala Phe Ile Leu Thr Arg Ala Glu Gly Tyr Pro Ser Val Phe Tyr Gly Asp Tyr Tyr Gly Thr Ser Gly Asn Ser Ser Tyr Glu Ile Pro Ala Leu Lys Asp Lys Ile Asp Pro Ile Leu Thr Ala Arg Lys Asn Phe Ala Tyr Gly Thr Gln Arg Asp Tyr Leu Asp His Pro Asp Val Ile Gly Trp Thr Arg Glu Gly Asp Ser Val His Ala Lys Ser Gly Leu Ala Ala Leu Ile Ser Asp Gly Pro Gly Gly Ser Lys Trp Met Asp Val Gly Lys Asn Asn Ala Gly Glu Val Trp Tyr Asp Ile Thr Gly Asn Gln Thr Asn Thr Val Thr Ile Asn Lys Asp Gly Ser Gly Gln Phe His Val Ser Gly Gly Ser Val Ser Ile Tyr Val Gln Gln

FIGURE 16HHH

SEQ ID NO: 147

atgagcttaaataacttaaggtaaaactgcttagttttgctgtgtcttctgctgtattgtcactggctccaaatttagccaatgctgcaaatttttgaaag
tgagatggtgataatccatccgtttcagtggacatatgacaatatagcaaaagagtgtacagagtaccttggtccagccggatttgacggtgtaca
gatttcccagccagcggaacataagcgggctgaaggagtatggtgggccgtatatcagccggttaattataagaattttacaaccatgaccggta
acgaggagcagcttaaggcaatgatcaagacctgtaatgatgcaggtgttaaggtgttcgctgacgctgttttcaaccaaaaggctacagacgg
tgtaggctggggcggttcaacttggagttataagaactaccctgacggattctccggatcagatttccatggagactgttccattgacaaaagctat
actgatgcaaataatgtcagaacctgtgcactctcaggtatgccggacgttgccacagataactccgctactcaggaaaagattgcagattacct
cgcttctttaatgaatatgggggtctatggtttccgtattgacgctgcaaagcacatgggatacaacgatatcaactccattctttcaaaaactgcac
agaagactggaagaagacctcctgcatatctggaagtaatcggagccggtaacgaagctgccgacattcagccggacaagtatacctttattga
gaatgcggttgtaactgacttcggttatgtctgggatgcaaatgagagtttcggaaagggtaattacggtaaggcactggaactcagtacctggct
cggtgcaaattcagaaacattcgtaaacaatcatgatgatgaatggggcagatgctcagccggtagctgctcaatgaaaactcagaattatgctg
attataatctggctcagtcctggcttgctgtatggcctgtaggtacagtaagacagatatattccggttattcattccctgtaaaagataatgatcctta
tcgcgtcagtgatgcaactcatgatcagggcgggcctcttggtgccgaccgctgtgaaggtggctggttgtgtcagcaccgtgtgtccttcgttct
caattccccaagatttgcgagagctaccagaggtactgctgtatcaaccaagggatttgacaatggtgctttgtggtttaacagaggaagcaaag
gtttttatgcacagaatactaccaacagtcctataacccagacattctctgttgaagtacctgacggaaattactgtgatatcttaggaacatcagat
cctaagagcaatccatgcggagcagacgttgtcgtaagcggcggtaaggctacctttactattcctgcaaagacagctgtggctatctgtacaga
ctcagactggtgcggcaagggggttgatccttgtgaaagtgatccgaccggtgctgcctgtgtttgtaaggggaaaccaccgttaatggtgtgt
gcgtcagctggtgtaatgcgcattcatcaaatgaggaatgcacctgtgtattgaatccgaatgatgccaactgtcaggctgatattgaacctacca
agggtaaactctgttacgccggtacttcaaacgggtggaaacaggatcctttaacatataaccgtaaaacaggtttctggactattaatctgactctt
gacggtgcaggtgataccagcggagctcagcgcttcaaggttacagacggatgttcatggaccggaacagtttacggttcttcaggtactgccg
gaaagttggatgtaaatacatcatcaaccggcgatgaacctgtgtctcttgttggtgattatgttctttccattaacgataagaccatggaatatacat
tcaccaaggcagatgaagtaactaatcagccaccggttgcatcatttaccgcgacagttaacggtctgaccgtttcttttgccaataattcatccga
ccctgagaatgatgaattaacctacagctggaatttcggtaatggtaaaacatcatccgagaaagctcctagcataacctatgaagaatccggta
agtatactgttactttaaaggttactgattcagctaataacactgatacatttactaaagatataactgtaacagcaccttctagtggcaagtacttaaa
ggttgcagtcagaggttcgcatgataattacggaactgatctgttaaccaagaacggttctgattggaccggcgtctttgaattctttggatccacta
gtgtcgacctgcaggcgcgcgagctc

SEQ ID NO: 148

Met Ser Leu Asn Asn Phe Lys Val Lys Leu Leu Ser Phe Ala Val Ser Ser Ala Val Leu Ser Leu Ala
Pro Asn Leu Ala Asn Ala Ala Asn Phe Glu Ser Glu Met Val Ile Ile His Pro Phe Gln Trp Thr Tyr
Asp Asn Ile Ala Lys Glu Cys Thr Glu Tyr Leu Gly Pro Ala Gly Phe Asp Gly Val Gln Ile Ser Gln
Pro Ala Glu His Lys Arg Ala Glu Gly Val Trp Trp Ala Val Tyr Gln Pro Val Asn Tyr Lys Asn Phe
Thr Thr Met Thr Gly Asn Glu Glu Gln Leu Lys Ala Met Ile Lys Thr Cys Asn Asp Ala Gly Val Lys
Val Phe Ala Asp Ala Val Phe Asn Gln Lys Ala Thr Asp Gly Val Gly Trp Gly Gly Ser Thr Trp Ser
Tyr Lys Asn Tyr Pro Asp Gly Phe Ser Gly Ser Asp Phe His Gly Asp Cys Ser Ile Asp Lys Ser Tyr
Thr Asp Ala Asn Asn Val Arg Thr Cys Ala Leu Ser Gly Met Pro Asp Val Ala Thr Asp Asn Ser Ala
Thr Gln Glu Lys Ile Ala Asp Tyr Leu Ala Ser Leu Met Asn Met Gly Val Tyr Gly Phe Arg Ile Asp
Ala Ala Lys His Met Gly Tyr Asn Asp Ile Asn Ser Ile Leu Ser Lys Thr Ala Gln Lys Thr Gly Arg
Arg Pro Pro Ala Tyr Leu Glu Val Ile Gly Ala Gly Asn Glu Ala Ala Asp Ile Gln Pro Asp Lys Tyr
Thr Phe Ile Glu Asn Ala Val Val Thr Asp Phe Gly Tyr Val Trp Asp Ala Asn Glu Ser Phe Gly Lys
Gly Asn Tyr Gly Lys Ala Leu Glu Leu Ser Thr Trp Leu Gly Ala Asn Ser Glu Thr Phe Val Asn Asn
His Asp Asp Glu Trp Gly Arg Cys Ser Ala Gly Ser Cys Ser Met Lys Thr Gln Asn Tyr Ala Asp Tyr
Asn Leu Ala Gln Ser Trp Leu Ala Val Trp Pro Val Gly Thr Val Arg Gln Ile Tyr Ser Gly Tyr Ser
Phe Pro Val Lys Asp Asn Asp Pro Tyr Arg Val Ser Asp Ala Thr His Asp Gln Gly Gly Pro Leu Gly
Ala Asp Arg Cys Glu Gly Gly Trp Leu Cys Gln His Arg Val Ser Phe Val Leu Asn Ser Pro Arg Phe
Ala Arg Ala Thr Arg Gly Thr Ala Val Ser Thr Lys Gly Phe Asp Asn Gly Ala Leu Trp Phe Asn Arg
Gly Ser Lys Gly Phe Tyr Ala Gln Asn Thr Thr Asn Ser Pro Ile Thr Gln Thr Phe Ser Val Glu Val
Pro Asp Gly Asn Tyr Cys Asp Ile Leu Gly Thr Ser Asp Pro Lys Ser Asn Pro Cys Gly Ala Asp Val
Val Val Ser Gly Gly Lys Ala Thr Phe Thr Ile Pro Ala Lys Thr Ala Val Ala Ile Cys Thr Asp Ser

FIGURE 16III

Asp Trp Cys Gly Lys Gly Val Asp Pro Cys Glu Ser Asp Pro Thr Gly Ala Ala Cys Val Cys Lys Gly
Glu Thr Thr Val Asn Gly Val Cys Val Ser Trp Cys Asn Ala His Ser Ser Asn Glu Glu Cys Thr Cys
Val Leu Asn Pro Asn Asp Ala Asn Cys Gln Ala Asp Ile Glu Pro Thr Lys Gly Lys Leu Cys Tyr Ala
Gly Thr Ser Asn Gly Trp Lys Gln Asp Pro Leu Thr Tyr Asn Arg Lys Thr Gly Phe Trp Thr Ile Asn
Leu Thr Leu Asp Gly Ala Gly Asp Thr Ser Gly Ala Gln Arg Phe Lys Val Thr Asp Gly Cys Ser Trp
Thr Gly Thr Val Tyr Gly Ser Ser Gly Thr Ala Gly Lys Leu Asp Val Asn Thr Ser Ser Thr Gly Asp
Glu Pro Val Ser Leu Val Gly Asp Tyr Val Leu Ser Ile Asn Asp Lys Thr Met Glu Tyr Thr Phe Thr
Lys Ala Asp Glu Val Thr Asn Gln Pro Pro Val Ala Ser Phe Thr Ala Thr Val Asn Gly Leu Thr Val
Ser Phe Ala Asn Asn Ser Ser Asp Pro Glu Asn Asp Glu Leu Thr Tyr Ser Trp Asn Phe Gly Asn Gly
Lys Thr Ser Ser Glu Lys Ala Pro Ser Ile Thr Tyr Glu Glu Ser Gly Lys Tyr Thr Val Thr Leu Lys
Val Thr Asp Ser Ala Asn Asn Thr Asp Thr Phe Thr Lys Asp Ile Thr Val Thr Ala Pro Ser Ser Gly
Lys Tyr Leu Lys Val Ala Val Arg Gly Ser His Asp Asn Tyr Gly Thr Asp Leu Leu Thr Lys Asn Gly
Ser Asp Trp Thr Gly Val Phe Glu Phe Phe Gly Ser Thr Ser Val Asp Leu Gln Ala Arg Glu Leu

SEQ ID NO: 149

atgatcttaagtaattttaaggtaaaacttcttagttttgctgtgtcttctgctgtactgacactggctgcaaatgtcgccaatgccaagaattatgaaa
gtgaaatggttattattcatccatttcagtggacatatgacaatatagcaaaagaatgtactgagtatctgggacctgcgggatttgacggggtgca
gatttcccaggcggctgagcataaagatgccggtggtgcatggtggggtacctaccagcctgtaaacttcaagagttttactaccatggttggta
atgaagaacagcttagagcaatgattaaaacctgtaacgaggcaggtgttaaggtctttgccgatgccgtgattaatcagaaagccggcgacgg
tgtaggtataggtggttcaactttcggaaattataattatcctgacggatttaccagtgatgattttcatcataataactgcagtataggtaataattatt
cagatgcatgggtagtaagattctgtgacctcagtggcatgccggatatagcaactgataacgacagtaccagaaataagattgctgattacttcg
ccagccttatgaatatgggggtatacggattccgtattgatgctgccaagcactttagctatgatgatatagacgctattgtagagaaaacagcaa
ccaaagcaggcaggagacctcctgtctatatggaggttatcggtaatccgggtcaagaggcggatgatatccagccgaacaagtatacatgga
ttgataatgccgttgtaacagattttacttatgctaatagcatgcataatatttttaacggaagcggttatgccaaggctttgaacatggggctagggc
atgttgatgctgaaaatgccgaagtctttataagtaatcatgataatgaatggggaagaaagtctgccggttcctgctcaataagaacccagaata
atccggattaccatctggctcagtcctggctcgcagtttggcctttaggcaaggttagacagatttattctgcatatcagttcccggtctttgaagata
gttgtgagcgggtcagtcagcaagcccatgatcagggcggtcctatcggggcagcccgctgtgaaggtggctggttgtgtcagcaccgtgtac
cgtttgtgctcaattctcctagatttgcaagagcaaccagagggacagtcgttactactaaaggttttgatgacggagctttgtggtttaacagagg
aagcaagggcttctatgcccagaatactaccggcagttctataactcatacattctcagttgaattacctgatggaaattactgtgatatccttggag
caaccgatccgaagaataatccttgcggagcggatgtcactgtaagcggaggtaaagcaacctttaccattccggcaaagaccgccgtagcta
tctgtactgatgaaaagtggtgtggcaaggggttgacccttgtgaaagcgatcctaccggttccgcctgtgtatgtaagggtgaaaccacagtt
aacggcgtatgtgtaagctggtgtaatgctcactcatctaatgaagaatgtgcctgtgtgctaaatcctaatgacgctgagtgtcaggccgacatt
gagccgaccaagggtaaactctgctatgtaggtacctccaacaagtggactcaggaacctttaacctataatcgcaagaccggtttctggactct
caacgttgaacttgacggtaaggggcgatacccagcggggcgcagcgctttaaagttaccgacggctgttcatggcagggtactgtttacggttca
tcaggagtagaaggcagacttgacgtaaatacttcagccaccggagatgaaccggtttcactgacaggtaaatatgttctttccataaatgataag
accatggaatacacattcattcctgcaggcagtggaaacaagcctccggttgcgtcatttactccgactgttaaagatctgactgtatcttttgtcaa
taattcatccgaccctgagaatgatgaattaacctacagctggaatttcggtaacggtaaaacctcatctgaaaagaatccgagtgttacatatgat
aaagccggtaaatatactgtttcactcaaagtaaaccgatactgcaaacaatcactgatacccaaaacactggaaatcgatttacatctcctgttaacg
gaaaatattccaaggttgcagtcagaggttcacatgataactacggaacaaatctgttaaccaggaatggttcagaatggaccggtatcttgaatt
cagtaagacaaccaaattcaagcttgaagctctgcctcctgcagctgaccagtgtatcttcctcggcggtaatcgaggtgaggcattgactgcct
ccggtggatttatatctcttcctgccggaaggtatactataaagtttaatgaggaaagcaaggttcttactgcaggcgatgttgactgcaccggg

SEQ ID NO: 150

Met Ile Leu Ser Asn Phe Lys Val Lys Leu Leu Ser Phe Ala Val Ser Ser Ala Val Leu Thr Leu Ala
Ala Asn Val Ala Asn Ala Lys Asn Tyr Glu Ser Glu Met Val Ile Ile His Pro Phe Gln Trp Thr Tyr
Asp Asn Ile Ala Lys Glu Cys Thr Glu Tyr Leu Gly Pro Ala Gly Phe Asp Gly Val Gln Ile Ser Gln
Ala Ala Glu His Lys Asp Ala Gly Gly Ala Trp Trp Gly Thr Tyr Gln Pro Val Asn Phe Lys Ser Phe
Thr Thr Met Val Gly Asn Glu Glu Gln Leu Arg Ala Met Ile Lys Thr Cys Asn Glu Ala Gly Val Lys
Val Phe Ala Asp Ala Val Ile Asn Gln Lys Ala Gly Asp Gly Val Gly Ile Gly Gly Ser Thr Phe Gly
Asn Tyr Asn Tyr Pro Asp Gly Phe Thr Ser Asp Asp Phe His His Asn Asn Cys Ser Ile Gly Asn Asn

FIGURE 16JJJ

Tyr Ser Asp Ala Trp Val Val Arg Phe Cys Asp Leu Ser Gly Met Pro Asp Ile Ala Thr Asp Asn Asp Ser Thr Arg Asn Lys Ile Ala Asp Tyr Phe Ala Ser Leu Met Asn Met Gly Val Tyr Gly Phe Arg Ile Asp Ala Ala Lys His Phe Ser Tyr Asp Asp Ile Asp Ala Ile Val Glu Lys Thr Ala Thr Lys Ala Gly Arg Arg Pro Pro Val Tyr Met Glu Val Ile Gly Asn Pro Gly Gln Glu Ala Asp Asp Ile Gln Pro Asn Lys Tyr Thr Trp Ile Asp Asn Ala Val Val Thr Asp Phe Thr Tyr Ala Asn Ser Met His Asn Ile Phe Asn Gly Ser Gly Tyr Ala Lys Ala Leu Asn Met Gly Leu Gly His Val Asp Ala Glu Asn Ala Glu Val Phe Ile Ser Asn His Asp Asn Glu Trp Gly Arg Lys Ser Ala Gly Ser Cys Ser Ile Arg Thr Gln Asn Asn Pro Asp Tyr His Leu Ala Gln Ser Trp Leu Ala Val Trp Pro Leu Gly Lys Val Arg Gln Ile Tyr Ser Ala Tyr Gln Phe Pro Val Phe Glu Asp Ser Cys Glu Arg Val Ser Gln Gln Ala His Asp Gln Gly Gly Pro Ile Gly Ala Ala Arg Cys Glu Gly Gly Trp Leu Cys Gln His Arg Val Pro Phe Val Leu Asn Ser Pro Arg Phe Ala Arg Ala Thr Arg Gly Thr Val Val Thr Thr Lys Gly Phe Asp Asp Gly Ala Leu Trp Phe Asn Arg Gly Ser Lys Gly Phe Tyr Ala Gln Asn Thr Thr Gly Ser Ser Ile Thr His Thr Phe Ser Val Glu Leu Pro Asp Gly Asn Tyr Cys Asp Ile Leu Gly Ala Thr Asp Pro Lys Asn Asn Pro Cys Gly Ala Asp Val Thr Val Ser Gly Gly Lys Ala Thr Phe Thr Ile Pro Ala Lys Thr Ala Val Ala Ile Cys Thr Asp Glu Lys Trp Cys Gly Lys Gly Val Asp Pro Cys Glu Ser Asp Pro Thr Gly Ser Ala Cys Val Cys Lys Gly Glu Thr Thr Val Asn Gly Val Cys Val Ser Trp Cys Asn Ala His Ser Ser Asn Glu Glu Cys Ala Cys Val Leu Asn Pro Asn Asp Ala Glu Cys Gln Ala Asp Ile Glu Pro Thr Lys Gly Lys Leu Cys Tyr Val Gly Thr Ser Asn Lys Trp Thr Gln Glu Pro Leu Thr Tyr Asn Arg Lys Thr Gly Phe Trp Thr Leu Asn Val Glu Leu Asp Gly Lys Gly Asp Thr Ser Gly Ala Gln Arg Phe Lys Val Thr Asp Gly Cys Ser Trp Gln Gly Thr Val Tyr Gly Ser Ser Gly Val Glu Gly Arg Leu Asp Val Asn Thr Ser Ala Thr Gly Asp Glu Pro Val Ser Leu Thr Gly Lys Tyr Val Leu Ser Ile Asn Asp Lys Thr Met Glu Tyr Thr Phe Ile Pro Ala Gly Ser Gly Asn Lys Pro Pro Val Ala Ser Phe Thr Pro Thr Val Lys Asp Leu Thr Val Ser Phe Val Asn Asn Ser Ser Asp Pro Glu Asn Asp Glu Leu Thr Tyr Ser Trp Asn Phe Gly Asn Gly Lys Thr Ser Ser Glu Lys Asn Pro Ser Val Thr Tyr Asp Lys Ala Gly Lys Tyr Thr Val Ser Leu Lys Val Thr Asp Thr Ala Asn Asn Thr Asp Thr Lys Thr Leu Glu Ile Asp Leu Thr Ser Pro Val Asn Gly Lys Tyr Ser Lys Val Ala Val Arg Gly Ser His Asp Asn Tyr Gly Thr Asn Leu Leu Thr Arg Asn Gly Ser Glu Trp Thr Gly Ile Phe Glu Phe Ser Lys Thr Thr Lys Phe Lys Leu Glu Ala Leu Pro Pro Ala Ala Asp Gln Cys Ile Phe Leu Gly Gly Asn Arg Gly Glu Ala Leu Thr Ala Ser Gly Gly Phe Ile Ser Leu Pro Ala Gly Arg Tyr Thr Ile Lys Phe Asn Glu Glu Ser Lys Val Leu Thr Ala Gly Asp Val Asp Cys Thr Gly

SEQ ID NO: 151

atgaaaactattctttcaacaatcatggtgatggcggctgcggctgccaccaccgtagaggctcaaggctggccggaaaactacggcggcgtc
atgttgcagggattctactgggattcctattcagccaccaagtggactaaactggaagcacaggctgacgagatctgcaactatttctcgctggta
tgggtaccacagtcggcctataccggcagcagtacctccatgggctacgacccgctgtattacttcgaccagcattcatcgttcggcaccgaag
agcagctacggtcgttcatcagtacctacaagcaagaaaggaactggcatcatagccgatgtagttgtcaatcaccgaaagaatgtctcaaactg
ggtcgattttccggccgagacctacaacggtgtaacctatcagatggtaaggcaccggacatcgttcgaacgatgcgtggcggaaaagaaggcca
cttgggcaaatcaaaacggctacagtctctcctcaatgccgacgaaggcgaaggctgggacggcatgcggcgaccttggacacaagttcgca
gaacgtgcagaaatcggttcttgcctacaccaaatatctggttgacgacttaggctataccggattccgctacgatatggtaaagggatttgacgg
atcgcatgtagccgactacaacaccaatgccggcgtgcagttctctgtcggcgaatattgggacggcactgcatcgaaagtttacagttggatca
acagcaccaaaaagagcgatgtgccgcagtcggcagccttcgacttcgctttccgatacacctgccgcgatgccgtcaacaacaagaactgg
gcgaacctgaagaacacttccggtatcagcgatgccgattacaggcgctattcggttacgtttgttgaaaatcacgatacggaataccgttcagct
acggcttcccaggatcccatcaagggtgatacggttgccctcaatgcctggatgctggctatgccgggcacaccttgtgttttcctgaaacattgg
accgactgcaaggaagagatcaagaatctcatcgaggcacgtcgcctggtcggtattcacaaccagagcacctatgccgaatggatgagcgg
tgcagcctacatcggacgtaccgtaacaggtacgaacggcaccttacgtgttctgtgcggctcttatcagtataatgtagccgccaactacattca
gattctctcaggcaaaaactataaatactacgtactcaacacgctcgaggctccctggatcgggaaaggttccggctcgtacaccgaaggtgaa
accgtaacggttccgctcatcgccatatcggccgatgccaatgccaagctggtatataccaccgacggcacagacccccaccgcaacctcaaca
gccgtaaccagcggaacggaactgaccatcacttcggacgccgtcctgaaggttggtctgctttccggcggcatcgtcaggaacatacagagc
cgtacattcaccttccaggctgcaaacacctccgagtattacacagccaccatgcacgtatgcaaccagtccggagctctcaatccgctgtttgc
ctatgttgggcaggaccggacaacgagcagattaacggcaactggccgggcaccaagctcaccgctaccattaccgaaaacaacctttacct

FIGURE 16KKK

ggtacacgcagtcgttccagattccgaagaacgtggactatgtcgtgaactttgttttcaccacaaccggcggcggtacgcagacagtggatgtt
accggcatgaaggccgatgtctggtacattattaacagtaccaagagcggcaacaagtacacggtaaccgacgttacctcacagtattcttcgtt
agaggccatctttgatgaagaaaactccggctccttccctgtctatgacctgcagggacgccgcgtcagcgaaattagaaacaggacaattatat
cttcagaacggaaagaagatactcatcagataaacagaggttccgaaccattctcctattatgaaaatcagacacttagtaatctcagcactgctg
ggtttgggggggcttgtacaccatcagctgctcctcgtcggg

SEQ ID NO: 152
Met Lys Thr Ile Leu Ser Thr Ile Met Val Met Ala Ala Ala Ala Ala Thr Thr Val Glu Ala Gln Gly
Trp Pro Glu Asn Tyr Gly Gly Val Met Leu Gln Gly Phe Tyr Trp Asp Ser Tyr Ser Ala Thr Lys Trp
Thr Lys Leu Glu Ala Gln Ala Asp Glu Ile Cys Asn Tyr Phe Ser Leu Val Trp Val Pro Gln Ser Ala
Tyr Thr Gly Ser Ser Thr Ser Met Gly Tyr Asp Pro Leu Tyr Tyr Phe Asp Gln His Ser Ser Phe Gly
Thr Glu Glu Gln Leu Arg Ser Phe Ile Ser Thr Tyr Lys Gln Lys Gly Thr Gly Ile Ile Ala Asp Val Val
Val Asn His Arg Lys Asn Val Ser Asn Trp Val Asp Phe Pro Ala Glu Thr Tyr Asn Gly Val Thr Tyr
Gln Met Val Ser Thr Asp Ile Val Ser Asn Asp Asp Gly Gly Lys Thr Ala Thr Trp Ala Asn Gln Asn
Gly Tyr Ser Leu Ser Ser Asn Ala Asp Glu Gly Glu Gly Trp Asp Gly Met Arg Asp Leu Asp His
Lys Ser Gln Asn Val Gln Lys Ser Val Leu Ala Tyr Thr Lys Tyr Leu Val Asp Asp Leu Gly Tyr Thr
Gly Phe Arg Tyr Asp Met Val Lys Gly Phe Asp Gly Ser His Val Ala Asp Tyr Asn Thr Asn Ala
Gly Val Gln Phe Ser Val Gly Glu Tyr Trp Asp Gly Thr Ala Ser Lys Val Tyr Ser Trp Ile Asn Ser
Thr Lys Lys Ser Asp Val Pro Gln Ser Ala Ala Phe Asp Phe Ala Phe Arg Tyr Thr Cys Arg Asp Ala
Val Asn Asn Lys Asn Trp Ala Asn Leu Lys Asn Thr Ser Gly Ile Ser Asp Ala Asp Tyr Arg Arg Tyr
Ser Val Thr Phe Val Glu Asn His Asp Thr Glu Tyr Arg Ser Ala Thr Ala Ser Gln Asp Pro Ile Lys
Gly Asp Thr Val Ala Leu Asn Ala Trp Met Leu Ala Met Pro Gly Thr Pro Cys Val Phe Leu Lys His
Trp Thr Asp Cys Lys Glu Glu Ile Lys Asn Leu Ile Glu Ala Arg Arg Leu Val Gly Ile His Asn Gln
Ser Thr Tyr Ala Glu Trp Met Ser Gly Ala Ala Tyr Ile Gly Arg Thr Val Thr Gly Thr Asn Gly Thr
Leu Arg Val Leu Cys Gly Ser Tyr Gln Tyr Asn Val Ala Ala Asn Tyr Ile Gln Ile Leu Ser Gly Lys
Asn Tyr Lys Tyr Tyr Val Leu Asn Thr Leu Glu Ala Pro Trp Ile Gly Lys Gly Ser Gly Ser Tyr Thr
Glu Gly Glu Thr Val Thr Val Pro Leu Ile Ala Ile Ser Ala Asp Ala Asn Ala Lys Leu Val Tyr Thr
Thr Asp Gly Thr Asp Pro Thr Ala Thr Ser Thr Ala Val Thr Ser Gly Thr Glu Leu Thr Ile Thr Ser
Asp Ala Val Leu Lys Val Gly Leu Leu Ser Gly Gly Ile Val Arg Asn Ile Gln Ser Arg Thr Phe Thr
Phe Gln Ala Ala Asn Thr Ser Glu Tyr Tyr Thr Ala Thr Met His Val Cys Asn Gln Ser Gly Ala Leu
Asn Pro Leu Phe Ala Tyr Val Trp Ala Gly Pro Asp Asn Glu Gln Ile Asn Gly Asn Trp Pro Gly Thr
Lys Leu Thr Ala Thr Ile Thr Glu Asn Asn Leu Thr Trp Tyr Thr Gln Ser Phe Gln Ile Pro Lys Asn
Val Asp Tyr Val Val Asn Phe Val Phe Thr Thr Thr Gly Gly Gly Thr Gln Thr Val Asp Val Thr Gly
Met Lys Ala Asp Val Trp Tyr Ile Ile Asn Ser Thr Lys Ser Gly Asn Lys Tyr Thr Val Thr Asp Val
Thr Ser Gln Tyr Ser Ser Leu Glu Ala Ile Phe Asp Glu Glu Asn Ser Gly Ser Phe Pro Val Tyr Asp
Leu Gln Gly Arg Arg Val Ser Glu Ile Arg Asn Arg Thr Ile Ile Ser Ser Glu Arg Lys Glu Asp Thr
His Gln Ile Asn Arg Gly Ser Glu Pro Phe Ser Tyr Tyr Glu Asn Gln Thr Leu Ser Asn Leu Ser Thr
Ala Gly Phe Gly Gly Leu Val His His Gln Leu Leu Leu Val Gly

SEQ ID NO: 69
atgttgaaaaggattacggtagtctgtttattgtttattttgctttttcctaatatatatgagggaaataaggcagaagcagcaacagtgaacaatgga
acattaatgcagtattttgagtggtacgctccgaatgatgggaatcattggaatcgtttgcgttccgatgctgaaagtttagctcataaaggaatcac
atctgtatggataccacctgcatataaagggacttcgcaaaatgatgtagggtatggggcctatgatttatatgatttaggggagttcaatcaaaaa
ggaacggtgcggacgaaatatgggacaaaagcacagttgaaatctgcaattgacgctttacataagcaaaacatcgacgtatacggtgatgtag
ttatgaatcataaaggtggggctgattatactgaaaccgtaacagctgttgaggtagaccgtaacaatcgaaatattgaagtatcaggtgattatca
aattagtgcatggacggggtttaattttccagggcgcggagatgcttattctaatttcaaatggaaatggtatcattttgacggaacggattgggatg
aaggaaggaaattaaatcgaatttataaatttaggggtgtagataaagcgtgggattgggaagtgtctagcgaaaatggaaattatgattatttgat
gtatgcagatcttgattttgatcatcctgatgttgcgaatgagatgaaaaattggggaacatggtatgcgaatgaattaaatttagatggctttcgttt
ggacgctgttaaacatattgatcatgaatatttacgcgattgggtaaatcatgccagacagcaaacggggaaagaaatgtttacagtagctgaata
ttggcaaaatgatgttcaggctttaaacaattatttagcgaaagtcaattataatcaatctgtgtttgatgcaccgcttcattacaatttcatttatgcttc

FIGURE 16LLL

aacaggaaatgggaattatgatatgagaaatattttaaatggaacagtaatgaaaaatcaccctgcactcgcagttactctcgttgagaatcatgat
tctcagcctgggcagtcattggaatctgtagtaagtccgtggtttaagccgctggcatatgcatttatttaactcgtgcagagggctatccttcagtt
ttctatggtgattactatgggacaagcggaaatagtagttatgaaattccagcgttaaaagataaaattgatccaattttgacggcacgaaaaaactt
tgcatatggtacgcagcgtgattatttagaccatccagatgtgattggctggacaagagaaggcgatggtgtacatgctaattctggtttagcgac
attactctcggacggaccaggaggatcaaagtggatggatgttggaaagaataacgctggggaagtatggtacgatattacgggtaatcaaac
aaatactgtaacaattaataaggacggatgggggcagttctatgtaagtggcggctcagtttccatatatgttcagcggtaa

SEQ ID NO: 70
Met Leu Lys Arg Ile Thr Val Val Cys Leu Leu Phe Ile Leu Leu Phe Pro Asn Ile Tyr Glu Gly Asn
Lys Ala Glu Ala Ala Thr Val Asn Asn Gly Thr Leu Met Gln Tyr Phe Glu Trp Tyr Ala Pro Asn Asp
Gly Asn His Trp Asn Arg Leu Arg Ser Asp Ala Glu Ser Leu Ala His Lys Gly Ile Thr Ser Val Trp
Ile Pro Pro Ala Tyr Lys Gly Thr Ser Gln Asn Asp Val Gly Tyr Gly Ala Tyr Asp Leu Tyr Asp Leu
Gly Glu Phe Asn Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly Thr Lys Ala Gln Leu Lys Ser Ala Ile
Asp Ala Leu His Lys Gln Asn Ile Asp Val Tyr Gly Asp Val Val Met Asn His Lys Gly Gly Ala Asp
Tyr Thr Glu Thr Val Thr Ala Val Glu Val Asp Arg Asn Asn Arg Asn Ile Glu Val Ser Gly Asp Tyr
Gln Ile Ser Ala Trp Thr Gly Phe Asn Phe Pro Gly Arg Gly Asp Ala Tyr Ser Asn Phe Lys Trp Lys
Trp Tyr His Phe Asp Gly Thr Asp Trp Asp Glu Gly Arg Lys Leu Asn Arg Ile Tyr Lys Phe Arg Gly
Val Asp Lys Ala Trp Asp Trp Glu Val Ser Ser Glu Asn Gly Asn Tyr Asp Tyr Leu Met Tyr Ala Asp
Leu Asp Phe Asp His Pro Asp Val Ala Asn Glu Met Lys Asn Trp Gly Thr Trp Tyr Ala Asn Glu
Leu Asn Leu Asp Gly Phe Arg Leu Asp Ala Val Lys His Ile Asp His Glu Tyr Leu Arg Asp Trp Val
Asn His Ala Arg Gln Gln Thr Gly Lys Glu Met Phe Thr Val Ala Glu Tyr Trp Gln Asn Asp Val Gln
Ala Leu Asn Asn Tyr Leu Ala Lys Val Asn Tyr Asn Gln Ser Val Phe Asp Ala Pro Leu His Tyr Asn
Phe His Tyr Ala Ser Thr Gly Asn Gly Asn Tyr Asp Met Arg Asn Ile Leu Asn Gly Thr Val Met Lys
Asn His Pro Ala Leu Ala Val Thr Leu Val Glu Asn His Asp Ser Gln Pro Gly Gln Ser Leu Glu Ser
Val Val Ser Pro Trp Phe Lys Pro Leu Ala Tyr Ala Phe Ile Leu Thr Arg Ala Glu Gly Tyr Pro Ser
Val Phe Tyr Gly Asp Tyr Tyr Gly Thr Ser Gly Asn Ser Ser Tyr Glu Ile Pro Ala Leu Lys Asp Lys
Ile Asp Pro Ile Leu Thr Ala Arg Lys Asn Phe Ala Tyr Gly Thr Gln Arg Asp Tyr Leu Asp His Pro
Asp Val Ile Gly Trp Thr Arg Glu Gly Asp Gly Val His Ala Asn Ser Gly Leu Ala Thr Leu Leu Ser
Asp Gly Pro Gly Gly Ser Lys Trp Met Asp Val Gly Lys Asn Asn Ala Gly Glu Val Trp Tyr Asp Ile
Thr Gly Asn Gln Thr Asn Thr Val Thr Ile Asn Lys Asp Gly Trp Gly Gln Phe Tyr Val Ser Gly Gly
Ser Val Ser Ile Tyr Val Gln Arg

SEQ ID NO: 153
ttgccttcaattaatgcaagcgattgcaaaaaaaagggagataggagtatgaagaggaaaaaatggactgcgttagcactatctttaccactagtt
atgagcttatcaacaaacatacaagcagaaacattacataataataagggtcaaaaggcgcaaacaggaaataaagacggaatttttttatgaact
gtatgttaattctttttatgatactgatagcaatggacatggtgatttaaaaggcgtcacaaagaaacttgattatttaaatgatggaaatccaagaac
aaatnatgaicttcaaatnaaaggtatctggatgatgcctattaacacctctcctagttatcaaaatatgatgtaacgattactatatatcgatcct
cagtatggaagtttacaagatttcgtgaactacacagaagcgcataaacgcaacgtaaaggtagtaatggatcttgttattaatcatacaagc
agtgagcatccttggtttgtcgatgcattaaaaatataaaacagtaagtatcgagattactatattgggctgataaaaatacagacttaaatgaaaa
aggccoatggggtcaacaagtatggcacaaagcgtcgaacggagagtatttctacgcaacgttctgggaagggatgccggacttaaactatga
caaccctaaagtaagagaagaaatgattaaaatcgggaaattttggctcaaacaaggagctgatggctttcgtctagatgcagccatgcacatctt
taaagggcaaacacctgaaggagcaaagaaaaatattgaatggtggaatgaattccgcgacgcgatgagagaaacgaatccaaatacgtatct
agttggtgaaatatgggatcaaccagaagtagttgctccgtattatcaatcgttagattctacatttaacttcgacttagcatataaaatcgttaattcc
gttaaaaatggtactgatcaaggggtagccgcggcagctgttgcaacggatgagttatataaaacatataatccaaataaaattgatggaacgttt
ttaacgaatcatgaccaaaatcgtgtaatgagtgagttaaatggtgatgtaaacaaagcaaaatcagcagcctctattctgttgacactcccctggta
atccgttcatttattatggcgaagaaatcggcatgacaggccaaaaaccagatgagttgattcgtgagcctttccgttggtatgaagatgataaag
aaggtcaaacgagctgggagactccagtatataacattgatcataatggtgtttcagttgaagcacaagataaacaaaaagcttctcttctaagcc
attatcgtaaaatgattcgtgttcgtcagcaacacgatgaacttgtcaaaggtaatttagaacctatttctgtcaataattcacaggttgttgcctataat
cgtacgtataaaaataaatcaattcaagtgtaccataatatttcagacaagccggttacattaactgtttcaaacaaaggaaaactgatttttctagt
gaattaggagcaaaaaaggaaaaatcaacattagtaattccagcgaatacgacagtgctagtaaagtaa

FIGURE 16MMM

SEQ ID NO: 154

Met Pro Ser Ile Asn Ala Ser Asp Cys Lys Lys Lys Gly Asp Arg Ser Met Lys Arg Lys Lys Trp Thr Ala Leu Ala Leu Ser Leu Pro Leu Val Met Ser Leu Ser Thr Asn Ile Gln Ala Glu Thr Leu His Asn Asn Lys Gly Gln Lys Ala Gln Thr Gly Asn Lys Asp Gly Ile Phe Tyr Glu Leu Tyr Val Asn Ser Phe Tyr Asp Thr Asp Ser Asn Gly His Gly Asp Leu Lys Gly Val Thr Lys Lys Leu Asp Tyr Leu Asn Asp Gly Asn Pro Arg Thr Asn Asn Asp Leu Gln Ile Asn Gly Ile Trp Met Met Pro Ile Asn Thr Ser Pro Ser Tyr His Lys Tyr Asp Val Thr Asp Tyr Tyr Asn Ile Asp Pro Gln Tyr Gly Ser Leu Gln Asp Phe Arg Glu Leu Thr Thr Glu Ala His Lys Arg Asn Val Lys Val Val Ile Asp Leu Val Ile Asn His Thr Ser Ser Glu His Pro Trp Phe Val Asp Ala Leu Lys Asn Lys Asn Ser Lys Tyr Arg Asp Tyr Tyr Ile Trp Ala Asp Lys Asn Thr Asp Leu Asn Glu Lys Gly Pro Trp Gly Gln Gln Val Trp His Lys Ala Ser Asn Gly Glu Tyr Phe Tyr Ala Thr Phe Trp Glu Gly Met Pro Asp Leu Asn Tyr Asp Asn Pro Lys Val Arg Glu Glu Met Ile Lys Ile Gly Lys Phe Trp Leu Lys Gln Gly Ala Asp Gly Phe Arg Leu Asp Ala Ala Met His Ile Phe Lys Gly Gln Thr Pro Glu Gly Ala Lys Lys Asn Ile Glu Trp Trp Asn Glu Phe Arg Asp Ala Met Arg Glu Thr Asn Pro Asn Thr Tyr Leu Val Gly Glu Ile Trp Asp Gln Pro Glu Val Val Ala Pro Tyr Tyr Gln Ser Leu Asp Ser Thr Phe Asn Phe Asp Leu Ala Tyr Lys Ile Val Asn Ser Val Lys Asn Gly Thr Asp Gln Gly Val Ala Ala Ala Ala Val Ala Thr Asp Glu Leu Tyr Lys Thr Tyr Asn Pro Asn Lys Ile Asp Gly Thr Phe Leu Thr Asn His Asp Gln Asn Arg Val Met Ser Glu Leu Asn Gly Asp Val Asn Lys Ala Lys Ser Ala Ala Ser Ile Leu Leu Thr Leu Pro Gly Asn Pro Phe Ile Tyr Tyr Gly Glu Glu Ile Gly Met Thr Gly Gln Lys Pro Asp Glu Leu Ile Arg Glu Pro Phe Arg Trp Tyr Glu Asp Asp Lys Glu Gly Gln Thr Ser Trp Glu Thr Pro Val Tyr Asn Ile Asp His Asn Gly Val Ser Val Glu Ala Gln Asp Lys Gln Lys Ala Ser Leu Leu Ser His Tyr Arg Lys Met Ile Arg Val Arg Gln Gln His Asp Glu Leu Val Lys Gly Asn Leu Glu Pro Ile Ser Val Asn Asn Ser Gln Val Val Ala Tyr Asn Arg Thr Tyr Lys Asn Lys Ser Ile Gln Val Tyr His Asn Ile Ser Asp Lys Pro Val Thr Leu Thr Val Ser Asn Lys Gly Lys Leu Ile Phe Ser Ser Glu Leu Gly Ala Lys Lys Glu Lys Ser Thr Leu Val Ile Pro Ala Asn Thr Thr Val Leu Val Lys

SEQ ID NO: 155

gtgtcaagaatgtttgcaaaacgattcaaaacctctttactgccgttattcgctggatttttattgctgtttcatttggttctggcaggaccaacggctg
cgaatgctgaaacggctaacaaatcaaatgagcttacagcaccgtcgatcaaaagcggaaccattcttcatgcttggaattggtcgttcaatacgt
taaaacacaatatgaaggatattcatgatgcaggatatacagcgattcagacgtctccgattaaccaagtcaaggaagggaaccaaggaaataa
aaacatgtcgaactggtactggctctatcagccgacatcgtaccaaattggcaaccgttacttaggtactgaacaagaatttaaagaaatgtgtgc
agccgctgaagaatatggcataaaggttattgttgacgcggtcatcaatcataccaccagtgactatgccgcgatttccaatgagattaagagtatt
ccaaactggacacatggaaacacacaaattaaaaactggtctgatcgatgggatgtcacgcagaatgcattgctcgggctgtatgactggaata
cacaaaatacacaagtacagtcctatttgaaacggttcttagaaagagcattgaatgacggggcagacggttttcgatttgatgccgccaaacata
tagagcttccggatgatggcagttacggcagtcaattttggccgaatatcacaaatacatctgcagagttccaatacggagaaatcctgcaggat
agtgcttcaagagatgcttcatatgcgaattatatgaatgtgacagcgctctaactatgggaattccataggtccgctttaaagaatcgtaatctggg
cgtgtcgaatatctccactatgcatcagatgtggtctgcggacaagctagtgacatgggtagaatcgcatgataagtatgccaatgatgatgaag
agtcgacatggatgagcgatgatgatatccgtttaggctgggcggtgatagcttctcgttcaggcagtacgcctcttttctttccagacctgaggg
aggcggaaatggtgtgagattcccggggaaaagccaaataggcgatcgcgggagtgctttatttgaagatcaggctatcactgcggtcaatag
atttcacaatgtgatggctggacagcctgaggaactctcgaacccaaatggaaacaaccagatatttatgaatcagcgcggctcacatggcgttg
tgctggcaaatgcaggttcatcctctgtttctatcaatacgccaacaaaattgcctgatggcaggtatgataataaagctggggcaggttcatttca
agtaaatgacggtaaactgacaggcacgatcaatgccaggtctgtggctgtgcttatcctgatgatattgcaaaagcgcctcatgttttccttgag
aattacaaaacaggtgtaacacattctttcaatgatcaactgacgattacactgcgtgcagatgcgaatacaacaaaagccgtttatcaaatcaata
atggaccagagacggcgtttaaggatggagatcaattcacaatcggaaaaggagatccatttggcaaaacatacaccatcatgttaaaaggaac
gaacagtgatggtgtaacgaggaccgaggaatacagttttgttaaaagagatccagcttcggccaaaaccatcggctatcaaaatccgaatcatt
ggagccaggtaaatgcttatatctataaacatgatgggggccgggca

SEQ ID NO: 156

FIGURE 16NNN

Val Ser Arg Met Phe Ala Lys Arg Phe Lys Thr Ser Leu Leu Pro Leu Phe Ala Gly Phe Leu Leu Leu Phe His Leu Val Leu Ala Gly Pro Thr Ala Ala Asn Ala Glu Thr Ala Asn Lys Ser Asn Glu Leu Thr Ala Pro Ser Ile Lys Ser Gly Thr Ile Leu His Ala Trp Asn Trp Ser Phe Asn Thr Leu Lys His Asn Met Lys Asp Ile His Asp Ala Gly Tyr Thr Ala Ile Gln Thr Ser Pro Ile Asn Gln Val Lys Glu Gly Asn Gln Gly Asn Lys Asn Met Ser Asn Trp Tyr Trp Leu Tyr Gln Pro Thr Ser Tyr Gln Ile Gly Asn Arg Tyr Leu Gly Thr Glu Gln Glu Phe Lys Glu Met Cys Ala Ala Ala Glu Glu Tyr Gly Ile Lys Val Ile Val Asp Ala Val Ile Asn His Thr Thr Ser Asp Tyr Ala Ala Ile Ser Asn Glu Ile Lys Ser Ile Pro Asn Trp Thr His Gly Asn Thr Gln Ile Lys Asn Trp Ser Asp Arg Trp Asp Val Thr Gln Asn Ala Leu Leu Gly Leu Tyr Asp Trp Asn Thr Gln Asn Thr Gln Val Gln Ser Tyr Leu Lys Arg Phe Leu Glu Arg Ala Leu Asn Asp Gly Ala Asp Gly Phe Arg Phe Asp Ala Ala Lys His Ile Glu Leu Pro Asp Asp Gly Ser Tyr Gly Ser Gln Phe Trp Pro Asn Ile Thr Asn Thr Ser Ala Glu Phe Gln Tyr Gly Glu Ile Leu Gln Asp Ser Ala Ser Arg Asp Ala Ser Tyr Ala Asn Tyr Met Asn Val Thr Ala Ser Asn Tyr Gly His Ser Ile Arg Ser Ala Leu Lys Asn Arg Asn Leu Gly Val Ser Asn Ile Ser His Tyr Ala Ser Asp Val Ser Ala Asp Lys Leu Val Thr Trp Val Glu Ser His Asp Thr Tyr Ala Asn Asp Asp Glu Glu Ser Thr Trp Met Ser Asp Asp Asp Ile Arg Leu Gly Trp Ala Val Ile Ala Ser Arg Ser Gly Ser Thr Pro Leu Phe Phe Ser Arg Pro Glu Gly Gly Gly Asn Gly Val Arg Phe Pro Gly Lys Ser Gln Ile Gly Asp Arg Gly Ser Ala Leu Phe Glu Asp Gln Ala Ile Thr Ala Val Asn Arg Phe His Asn Val Met Ala Gly Gln Pro Glu Glu Leu Ser Asn Pro Asn Gly Asn Asn Gln Ile Phe Met Asn Gln Arg Gly Ser His Gly Val Val Leu Ala Asn Ala Gly Ser Ser Ser Val Ser Ile Asn Thr Pro Thr Lys Leu Pro Asp Gly Arg Tyr Asp Asn Lys Ala Gly Ala Gly Ser Phe Gln Val Asn Asp Gly Lys Leu Thr Gly Thr Ile Asn Ala Arg Ser Val Ala Val Leu Tyr Pro Asp Asp Ile Ala Lys Ala Pro His Val Phe Leu Glu Asn Tyr Lys Thr Gly Val Thr His Ser Phe Asn Asp Gln Leu Thr Ile Thr Leu Arg Ala Asp Ala Asn Thr Thr Lys Ala Val Tyr Gln Ile Asn Asn Gly Pro Glu Thr Ala Phe Lys Asp Gly Asp Gln Phe Thr Ile Gly Lys Gly Asp Pro Phe Gly Lys Thr Tyr Thr Ile Met Leu Lys Gly Thr Asn Ser Asp Gly Val Thr Arg Thr Glu Glu Tyr Ser Phe Val Lys Arg Asp Pro Ala Ser Ala Lys Thr Ile Gly Tyr Gln Asn Pro Asn His Trp Ser Gln Val Asn Ala Tyr Ile Tyr Lys His Asp Gly Gly Arg Ala

SEQ ID NO: 157

atgcaaacgattgcaaaaaaaggggatgaaacgatgaaagggaaaaaatggacagcattagctctaacactgccgctggctgctagcttatca
acaggcgttcacgccgaaaccgtacataaaggtaaagctccaacagcagataaaaacggtgtcttttatgaggtgtatgtaaactctttttacgat
gcaaataaagatggacatggtgatttaaaaggtcttacacaaaagctggattatttgaatgacggcaattctcataccaaaaatgatcttcaagtaa
acggaatttggatgatgccggtaaacccttctcctagctatcataaatatgatgtaacggactattataacattgatccgcagtacggaaatctgca
agattttcgcaagctgatgaaagaagcagataaacgagacgtaaaggttattatggacctcgttgtgaatcatacaagcagtgaacatccttggtt
tcaagctgcattaaaagataaaaacagcaagtacagagattactatatttgggccgataaaaatactgatttaaatgaaaaaggatcttgggggca
gcaagtatggcataaagctccaaacggagagtatttttatggtacgttttgggaaggaatgcctgacttaaattacgataatcccgaagtaagaaa
agaaatgattaacgtcgggaaattttggctaaagcaaggcgttgacgggttccgcttagatgctgcgcttcatatttttaaaggtcaaacacctgaa
ggcgctaagaaaaatatggtgtcgtggaatgagtttagtggatgcaatgaaaaaagaaaaccctaacgtatatctaacgggtgaagtatgggatc
aaccggtaagtcgagctcctactatcaatcgcttgattcttattaacttgatttagcaggaaagattgtaaactctgtaaaatcaggaaatgatca
aggaatcgcgactgcagcagccgcaactgatgagctgttcaaatcatacaatccaaataaattgacggcattttcttaaccaaccatgaccaaaa
atcgcgtcatgagtgagctaagcggcgatgtgaataaagcaaagtcagctgcctctatcttacttacgcttcctggcaacccgtatatttattacgg
tgaagaaattggaatgaccggtgaaaagcctgatgagttaatccgtgaaccgttccgctggtacgaaggcaatggacttggacaaaccagctg
ggaaacatccgtatacaacaaaggcggcaatggtgtgtcagtagagacacaaacaaaacaaaaggattctttgttaaatcattaccgtgaaatga
ttcgcgtgcgtcagcagcatgaagagttagtaaaaggaacccttcaatctatttcagtagacagtaaagaagtcgttgcctatagccgcacgtata
aaggcaaatcgattagcgtgtatcataatatttcaaatcaaccggtaaaagtatctgtaacagcgaaaggtaaattgatttttgctagtgaaaaaggt
gcaaaaaaagtcaaaaatcagcttgtggttccagctaatacaacggttttaataaaataa

SEQ ID NO: 158

Met Gln Thr Ile Ala Lys Lys Gly Asp Glu Thr Met Lys Gly Lys Lys Trp Thr Ala Leu Ala Leu Thr Leu Pro Leu Ala Ala Ser Leu Ser Thr Gly Val His Ala Glu Thr Val His Lys Gly Lys Ala Pro Thr Ala Asp Lys Asn Gly Val Phe Tyr Glu Val Tyr Val Asn Ser Phe Tyr Asp Ala Asn Lys Asp Gly His

FIGURE 16000

Gly Asp Leu Lys Gly Leu Thr Gln Lys Leu Asp Tyr Leu Asn Asp Gly Asn Ser His Thr Lys Asn Asp Leu Gln Val Asn Gly Ile Trp Met Met Pro Val Asn Pro Ser Pro Ser Tyr His Lys Tyr Asp Val Thr Asp Tyr Tyr Asn Ile Asp Pro Gln Tyr Gly Asn Leu Gln Asp Phe Arg Lys Leu Met Lys Glu Ala Asp Lys Arg Asp Val Lys Val Ile Met Asp Leu Val Val Asn His Thr Ser Ser Glu His Pro Trp Phe Gln Ala Ala Leu Lys Asp Lys Asn Ser Lys Tyr Arg Asp Tyr Tyr Ile Trp Ala Asp Lys Asn Thr Asp Leu Asn Glu Lys Gly Ser Trp Gly Gln Gln Val Trp His Lys Ala Pro Asn Gly Glu Tyr Phe Tyr Gly Thr Phe Trp Glu Gly Met Pro Asp Leu Asn Tyr Asp Asn Pro Glu Val Arg Lys Glu Met Ile Asn Val Gly Lys Phe Trp Leu Lys Gln Gly Val Asp Gly Phe Arg Leu Asp Ala Ala Leu His Ile Phe Lys Gly Gln Thr Pro Glu Gly Ala Lys Lys Asn Ile Val Trp Trp Asn Glu Phe Arg Asp Ala Met Lys Lys Glu Asn Pro Asn Val Tyr Leu Thr Gly Glu Val Trp Asp Gln Pro Glu Val Val Ala Pro Tyr Tyr Gln Ser Leu Asp Ser Leu Phe Asn Phe Asp Leu Ala Gly Lys Ile Val Asn Ser Val Lys Ser Gly Asn Asp Gln Gly Ile Ala Thr Ala Ala Ala Ala Thr Asp Glu Leu Phe Lys Ser Tyr Asn Pro Asn Lys Ile Asp Gly Ile Phe Leu Thr Asn His Asp Gln Asn Arg Val Met Ser Glu Leu Ser Gly Asp Val Asn Lys Ala Lys Ser Ala Ala Ser Ile Leu Leu Thr Leu Pro Gly Asn Pro Tyr Ile Tyr Tyr Gly Glu Glu Ile Gly Met Thr Gly Glu Lys Pro Asp Glu Leu Ile Arg Glu Pro Phe Arg Trp Tyr Glu Gly Asn Gly Leu Gly Gln Thr Ser Trp Glu Thr Ser Val Tyr Asn Lys Gly Gly Asn Gly Val Ser Val Glu Thr Gln Thr Lys Gln Lys Asp Ser Leu Leu Asn His Tyr Arg Glu Met Ile Arg Val Arg Gln Gln His Glu Glu Leu Val Lys Gly Thr Leu Gln Ser Ile Ser Val Asp Ser Lys Glu Val Val Ala Tyr Ser Arg Thr Tyr Lys Gly Lys Ser Ile Ser Val Tyr His Asn Ile Ser Asn Gln Pro Val Lys Val Ser Val Thr Ala Lys Gly Lys Leu Ile Phe Ala Ser Glu Lys Gly Ala Lys Lys Val Lys Asn Gln Leu Val Val Pro Ala Asn Thr Thr Val Leu Ile Lys

SEQ ID NO: 159

ttgcaaaaaaaaggggatgaaacgatgaaagggaaaaaatggacagctttagctctaacactgccgctggctgctagcttatcaacaggcgttc
acgccgaaaccgtacataaaggtaaatctccaacagcagataaaaacggtgtatttatgaggtgtatgtaaactctttttacgatgcaaataaaga
tggacatggtgatttaaaaggtcttacacaaaagttggattatttaaatgatggcaattctcatacaaagaatgatcttcaagtaaacgggatttggat
gatgccggtcaacccttctcccagctatcataaatatgatgtaacggactattataatattgatccgcagtatggaaatctgcaagattttcgcaaac
tgatgaaagaagcagataaacgagatgtaaaagtcattatggacctcgttgtgaatcatacgagcagtgaacacccttggtttcaagctgcattaa
aagataaaaacagcaagtacagagattactatatctgggctgataaaaataccgacttgaatgaaaaaggatcttggggacagcaagtatggca
taaagctccaaacggagagtatttttacggaacgttttgggaaggaatgccggacttaaattacgataatcctgaagtaagaaaagaaatgattaa
cgtaggaaagttttggctaaagcaaggagttgatgggttccgtctagatgctgcgcttcatatttttaaaggccaaacacctgaaggcgctaagaa
aaatctcctgtggtggaatgaatttagagatgcaatgaaaaaggaaaaccctaacgtatatctaacgggtgaagtatgggatcaaccggaagta
gtagctccttactatcaatcgcttgattctttatttaactttgatttagcaggaaagattgtaaactctgtaaaatcaggaaatgatcaaggaatcgcga
ctgcagcagcggcaacggatgaactgttcaaatcatacaatccaaataaaattgacggtatttttcttaaccaaccatgaccaaaatcgcgtcatga
gtgagctaaacggcgatgtgaataaagcaaagtcagctgcctctatcttacttacgcttcctggcaacccgtatatttattacggtgaagaaatcgg
catgaccggtgaaaagcctgatgagttaatccgtgaaccgttcccctggtacgaaggaaacggacttggacaaaccagctgggaaacacctgt
atataacaaaggcggcaacggcgtgtctgtagaagcacaaacaaaacaaaaggactctttgttaaatcattaccgtgaaatgattcgcgtgcgtc
agcagcacgaagagttagtaaaaggaacggttcaatctatttcagtagcacagtaaagaagtcgttgcctatagcgtacgtataaaggcaaatcg
attagcgtgtatcataatatttcaaatcaaccggtaaaagtatctgtagcagcaaaaggtaaattgattttgctagtgaaaaaggtgctaagaaagt
caaaaatcagcttgtgattccggcgaatacaacggttttaataaaataa

SEQ ID NO: 160

Met Gln Lys Lys Gly Asp Glu Thr Met Lys Gly Lys Lys Trp Thr Ala Leu Ala Leu Thr Leu Pro Leu Ala Ala Ser Leu Ser Thr Gly Val His Ala Glu Thr Val His Lys Gly Lys Ser Pro Thr Ala Asp Lys Asn Gly Val Phe Tyr Glu Val Tyr Val Asn Ser Phe Tyr Asp Ala Asn Lys Asp Gly His Gly Asp Leu Lys Gly Leu Thr Gln Lys Leu Asp Tyr Leu Asn Asp Gly Asn Ser His Thr Lys Asn Asp Leu Gln Val Asn Gly Ile Trp Met Met Pro Val Asn Pro Ser Pro Ser Tyr His Lys Tyr Asp Val Thr Asp Tyr Tyr Asn Ile Asp Pro Gln Tyr Gly Asn Leu Gln Asp Phe Arg Lys Leu Met Lys Glu Ala Asp Lys Arg Asp Val Lys Val Ile Met Asp Leu Val Val Asn His Thr Ser Ser Glu His Pro Trp Phe Gln Ala Ala Leu Lys Asp Lys Asn Ser Lys Tyr Arg Asp Tyr Tyr Ile Trp Ala Asp Lys Asn Thr Asp Leu Asn Glu Lys Gly Ser Trp Gly Gln Gln Val Trp His Lys Ala Pro Asn Gly Glu Tyr Phe Tyr Gly Thr

FIGURE 16PPP

Phe Trp Glu Gly Met Pro Asp Leu Asn Tyr Asp Asn Pro Glu Val Arg Lys Glu Met Ile Asn Val Gly Lys Phe Trp Leu Lys Gln Gly Val Asp Gly Phe Arg Leu Asp Ala Ala Leu His Ile Phe Lys Gly Gln Thr Pro Glu Gly Ala Lys Lys Asn Leu Leu Trp Trp Asn Glu Phe Arg Asp Ala Met Lys Lys Glu Asn Pro Asn Val Tyr Leu Thr Gly Glu Val Trp Asp Gln Pro Glu Val Val Ala Pro Tyr Tyr Gln Ser Leu Asp Ser Leu Phe Asn Phe Asp Leu Ala Gly Lys Ile Val Asn Ser Val Lys Ser Gly Asn Asp Gln Gly Ile Ala Thr Ala Ala Ala Ala Thr Asp Glu Leu Phe Lys Ser Tyr Asn Pro Asn Lys Ile Asp Gly Ile Phe Leu Thr Asn His Asp Gln Asn Arg Val Met Ser Glu Leu Asn Gly Asp Val Asn Lys Ala Lys Ser Ala Ala Ser Ile Leu Leu Thr Leu Pro Gly Asn Pro Tyr Ile Tyr Tyr Gly Glu Glu Ile Gly Met Thr Gly Glu Lys Pro Asp Glu Leu Ile Arg Glu Pro Phe Pro Trp Tyr Glu Gly Asn Gly Leu Gly Gln Thr Ser Trp Glu Thr Pro Val Tyr Asn Lys Gly Gly Asn Gly Val Ser Val Glu Ala Gln Thr Lys Gln Lys Asp Ser Leu Leu Asn His Tyr Arg Glu Met Ile Arg Val Arg Gln Gln His Glu Glu Leu Val Lys Gly Thr Leu Gln Ser Ile Ser Val Asp Ser Lys Glu Val Val Ala Tyr Ser Arg Thr Tyr Lys Gly Lys Ser Ile Ser Val Tyr His Asn Ile Ser Asn Gln Pro Val Lys Val Ser Val Ala Ala Lys Gly Lys Leu Ile Phe Ala Ser Glu Lys Gly Ala Lys Lys Val Lys Asn Gln Leu Val Ile Pro Ala Asn Thr Thr Val Leu Ile Lys

SEQ ID NO: 161

gtggatccaaagaattgtagtcaatttatgcaaacgattgcaaaaaaaggggatgaaacgatgaaagggaaaaaatggacagctttagctctaa
cactgccgctggctgctagcttatcaacaggtgttcacgccgaaaccgtacataaaggtaaagctccaacagcagataaaaacggtgtcttttat
gaggtatatgtaaactctttttacgatgcaaataaagatggacatggtgatttaaaaggccttacacaaaagttggactatttaaatgacggaaattc
tcatacaaagaatgatcttcaagtaaacgggatttggatgatgccggtcaacccttctcctagctatcataaatatgatgtaacggactattataatat
tgatccgcagtatggaaatctgcaagattttcgcaaacttatgaaagaagcagataaacgagacgtaaaagtcattatggaccttgttgtgaatcat
acgagcagtgaacacccttggtttcaagctgcgttgaaagataaaaacagcaagtacagagattactatatttgggctgataaaaatactgacttg
aatgaaaaaggatcttggggacaacaagtatggcataaagctccaaacggagagtatttttacggaacgttctgggaaggaatgcctgacttaa
attacgataaccctgaagtaagaaaagaaatgattaacgtcggaaagttttggctaaaacaaggcgttgacggcttccgcttagatgctgcccttc
atatttttaaaggtcaaacgcctgaaggcgctaagaaaaacattctatggtggaatgagtttagagatgcgatgaaaaaagaaaacccgaacgta
tatctaacgggtgaagtgtgggaccagccagaagtagtagccccttactatcaatcacttgattctctatttaattttgatttagcaggaaaaattgtc
agctctgtaaaagcaggaaatgatcaaggaatcgccactgcagcagcggcaactgatgagctgttcaaatcatacaatccaaataaaattgacg
gcattttcttaaccaaccatgaccaaaatcgcgtcatgagtgagttaagcggcgatgtgaataaagcaaaatcagccgcctctatcttacttacgct
tcctggaaatccgtatatttattacggtgaagaaattggcatgacaggtgaaaagcctgatgaattaatccgtgaaccgttccgctggtacgaagg
caacggaattggacaaactagctgggaaacacctgtatataacaaaggcggtaacggcgtgtctgtagaagcacaaacaaaacaaaaggatt
ccttgttaaatcattaccgtgaaatgattcgtgtgcgccagcagcacgaagagttagtaaaaggaacgcttcaatccatttcagtagacagtaaag
aagtcgttgcctatagccgcacgtacaaaggcaaatcgattagcgtgtatcataatatttcaaatcaacctgtaaaagtatctgtagcagcgaaag
gtaacttgatttttgctagtgaaaaaggtgctaagaaagtcaaaaatcagcttgtgattccggcgaatgcgacggttttaataaaataa

SEQ ID NO: 162

Val Asp Pro Lys Asn Cys Ser Gln Phe Met Gln Thr Ile Ala Lys Lys Gly Asp Glu Thr Met Lys Gly Lys Lys Trp Thr Ala Leu Ala Leu Thr Leu Pro Leu Ala Ala Ser Leu Ser Thr Gly Val His Ala Glu Thr Val His Lys Gly Lys Ala Pro Thr Ala Asp Lys Asn Gly Val Phe Tyr Glu Val Tyr Val Asn Ser Phe Tyr Asp Ala Asn Lys Asp Gly His Gly Asp Leu Lys Gly Leu Thr Gln Lys Leu Asp Tyr Leu Asn Asp Gly Asn Ser His Thr Lys Asn Asp Leu Gln Val Asn Gly Ile Trp Met Met Pro Val Asn Pro Ser Pro Ser Tyr His Lys Tyr Asp Val Thr Asp Tyr Tyr Asn Ile Asp Pro Gln Tyr Gly Asn Leu Gln Asp Phe Arg Lys Leu Met Lys Glu Ala Asp Lys Arg Asp Val Lys Val Ile Met Asp Leu Val Val Asn His Thr Ser Ser Glu His Pro Trp Phe Gln Ala Ala Leu Lys Asp Lys Asn Ser Lys Tyr Arg Asp Tyr Tyr Ile Trp Ala Asp Lys Asn Thr Asp Leu Asn Glu Lys Gly Ser Trp Gly Gln Gln Val Trp His Lys Ala Pro Asn Gly Glu Tyr Phe Tyr Gly Thr Phe Trp Glu Gly Met Pro Asp Leu Asn Tyr Asp Asn Pro Glu Val Arg Lys Glu Met Ile Asn Val Gly Lys Phe Trp Leu Lys Gln Gly Val Asp Gly Phe Arg Leu Asp Ala Ala Leu His Ile Phe Lys Gly Gln Thr Pro Glu Gly Ala Lys Lys Asn Ile Leu Trp Trp Asn Glu Phe Arg Asp Ala Met Lys Lys Glu Asn Pro Asn Val Tyr Leu Thr Gly Glu Val Trp Asp Gln Pro Glu Val Val Ala Pro Tyr Tyr Gln Ser Leu Asp Ser Leu Phe Asn Phe Asp Leu Ala Gly Lys Ile Val Ser Ser Val Lys Ala Gly Asn Asp Gln Gly Ile Ala Thr Ala Ala Ala Ala Thr Asp Glu

FIGURE 16QQQ

Leu Phe Lys Ser Tyr Asn Pro Asn Lys Ile Asp Gly Ile Phe Leu Thr Asn His Asp Gln Asn Arg Val Met Ser Glu Leu Ser Gly Asp Val Asn Lys Ala Lys Ser Ala Ala Ser Ile Leu Leu Thr Leu Pro Gly Asn Pro Tyr Ile Tyr Tyr Gly Glu Glu Ile Gly Met Thr Gly Glu Lys Pro Asp Glu Leu Ile Arg Glu Pro Phe Arg Trp Tyr Glu Gly Asn Gly Ile Gly Gln Thr Ser Trp Glu Thr Pro Val Tyr Asn Lys Gly Gly Asn Gly Val Ser Val Glu Ala Gln Thr Lys Gln Lys Asp Ser Leu Leu Asn His Tyr Arg Glu Met Ile Arg Val Arg Gln Gln His Glu Glu Leu Val Lys Gly Thr Leu Gln Ser Ile Ser Val Asp Ser Lys Glu Val Val Ala Tyr Ser Arg Thr Tyr Lys Gly Lys Ser Ile Ser Val Tyr His Asn Ile Ser Asn Gln Pro Val Lys Val Ser Val Ala Ala Lys Gly Asn Leu Ile Phe Ala Ser Glu Lys Gly Ala Lys Lys Val Lys Asn Gln Leu Val Ile Pro Ala Asn Ala Thr Val Leu Ile Lys

SEQ ID NO: 163

atggtacgtcccgaacgacgggctgcattggaaccgactatcgaacgactcgcagcacttgaaagacattgggtgacgacggtgtggattccg
ccggcgtacaaaggcacgtcacagaacgatgtcgggtatggggcgtacgatttatacgatctcggcgaattcaaccaaaaagggacgacccg
gacgaagtacgggacgaaagcgcagctccagaccgccatctcgaacttgcgcggtaaagggatcggtgtgtacggcgacgtcgtcatgaat
cacaagggcggggccgattataccgaatccgttcaggcgatcgaggtcaatccgtcgaaccggaaccaagaaacgtccggtgagtatggcat
ctcggcctggactgggttcaacttcgcggggcgcaacaatacatactcgccgttcaaatggcgctggtaccattttgacggtaccgattgggatc
agtcacgcagcttgagccgcatctataagttcaagagcacaggcaaggcgtgggacacggacgtgtcgaacgagaacggcaactatgattat
cttatgtatgccgacgtcgatttcgagcatcccgaggtccgccaagagatgaagaactggggcaaatggtacgccgactcgctcgggctcgac
ggtttccggttggatgcggtcaaacatatcagccactcgtacttgaaggagtgggtgacgagcgtgcgccagacgaccgggaaagagatgttc
acggtcgccgagtattggaagaacgatctcggtgccatcaacgactatctgtataagacgggctacacgcactccgtcttcgatgtgccgctcc
attataacttccaagcggccggtaacggcggcgggtattacgatatgcgcaacatcttgaaaggcaccgtcaccgaacagcatccgtcgctgtc
cgtgacgattgtcgataaccacgactcacagccgggccagtcgctcgagtcgacggtcgccaactggttcaaaccgctcgcctacgcgacga
tcatgacgcgcggtcagggttatccggccctcttctatggagactattatggcacgaaagggacgacgaaccgcgaaatcccgaacatgtcgg
gcacgctccaaccgattttgaaggcacgaaaagacttcgcctacgggacgcagcatgactacctcgatcatcaggacgtcatcggctggacac
gtgaaggtgtgaccgaccgtgccaaatcgggtctcgcgacgattctatcggacggtccgggcggctcgaagtggatgtacgtcggcaaacag
aacgccggcgaggtatggaaagacatgacgaacaacaacgcccgtctcgtcacgatcaatgctgacggctggggtcagttcttcgtcaacgg
aggctcggtctcgatttatacgcaacaataa

SEQ ID NO: 164

Met Val Arg Pro Glu Arg Arg Ala Ala Leu Glu Pro Thr Ile Glu Arg Leu Ala Ala Leu Glu Arg His Trp Val Thr Thr Val Trp Ile Pro Pro Ala Tyr Lys Gly Thr Ser Gln Asn Asp Val Gly Tyr Gly Ala Tyr Asp Leu Tyr Asp Leu Gly Glu Phe Asn Gln Lys Gly Thr Thr Arg Thr Lys Tyr Gly Thr Lys Ala Gln Leu Gln Thr Ala Ile Ser Asn Leu Arg Gly Lys Gly Ile Gly Val Tyr Gly Asp Val Val Met Asn His Lys Gly Gly Ala Asp Tyr Thr Glu Ser Val Gln Ala Ile Glu Val Asn Pro Ser Asn Arg Asn Gln Glu Thr Ser Gly Glu Tyr Gly Ile Ser Ala Trp Thr Gly Phe Asn Phe Ala Gly Arg Asn Asn Thr Tyr Ser Pro Phe Lys Trp Arg Trp Tyr His Phe Asp Gly Thr Asp Trp Asp Gln Ser Arg Ser Leu Ser Arg Ile Tyr Lys Phe Lys Ser Thr Gly Lys Ala Trp Asp Thr Asp Val Ser Asn Glu Asn Gly Asn Tyr Asp Tyr Leu Met Tyr Ala Asp Val Asp Phe Glu His Pro Glu Val Arg Gln Glu Met Lys Asn Trp Gly Lys Trp Tyr Ala Asp Ser Leu Gly Leu Asp Gly Phe Arg Leu Asp Ala Val Lys His Ile Ser His Ser Tyr Leu Lys Glu Trp Val Thr Ser Val Arg Gln Thr Thr Gly Lys Glu Met Phe Thr Val Ala Glu Tyr Trp Lys Asn Asp Leu Gly Ala Ile Asn Asp Tyr Leu Tyr Lys Thr Gly Tyr Thr His Ser Val Phe Asp Val Pro Leu His Tyr Asn Phe Gln Ala Ala Gly Asn Gly Gly Gly Tyr Tyr Asp Met Arg Asn Ile Leu Lys Gly Thr Val Thr Glu Gln His Pro Ser Leu Ser Val Thr Ile Val Asp Asn His Asp Ser Gln Pro Gly Gln Ser Leu Glu Ser Thr Val Ala Asn Trp Phe Lys Pro Leu Ala Tyr Ala Thr Ile Met Thr Arg Gly Gln Gly Tyr Pro Ala Leu Phe Tyr Gly Asp Tyr Tyr Gly Thr Lys Gly Thr Thr Asn Arg Glu Ile Pro Asn Met Ser Gly Thr Leu Gln Pro Ile Leu Lys Ala Arg Lys Asp Phe Ala Tyr Gly Thr Gln His Asp Tyr Leu Asp His Gln Asp Val Ile Gly Trp Thr Arg Glu Gly Val Thr Asp Arg Ala Lys Ser Gly Leu Ala Thr Ile Leu Ser Asp Gly Pro Gly Gly Ser Lys Trp Met Tyr Val Gly Lys Gln Asn Ala Gly Glu Val Trp Lys Asp Met Thr Asn Asn Asn Ala Arg Leu Val Thr Ile Asn Ala Asp Gly Trp Gly Gln Phe Phe Val Asn Gly Gly Ser Val Ser Ile Tyr Thr Gln Gln

FIGURE 16RRR

SEQ ID NO: 165

atgcagtatttcgagtggtacgtgccaaatgatggggaacattggaatcgtttgcgtaatgatgctgaaaatttagctcataaaggaattacatctgt
atggataccacccgtatataaaggaacttcacaaaatgatgtagggtatggagtgtatgatgtatatgatttgggagaattcaatcaaaaaggaac
gatacggacaaaatatgggacaaaagcacaattaaaatctgcaattgaggctttacataatcaaaatatcgatgtatacggtgatgttgttatgaac
cataaaggtggggcagattatactgaggttgtaacagccgttgaggtagaccgtaacaatcgaaatattgaaacatcgagtgattatcaaatagat
gcgtggacgggatttgattttccaggacgcagggactcctattctaattttaaatggagatggtttcattttgatggaacagattgggatgagggaa
ggaaattaaatagaatttataaatttaaaggcgtaggtaaagcttgggactgggaagtgtctagtgagaatggtaactatgattatttaatgtatgca
gatcttgatttcgatcatcctgaagttgcaaatgaaatgaaaaactggggaacctggtatgcggacgaattaaatttagatggctttcgtttagacg
cagttaaacatattgaccatgagtatcttcgtgattgggtaaatcatgttagaaagcaaacggggaaggaaatgtttacagtagctgaatattggca
aaatgatattcgtactttaaacaattatttagggaaagtaaattataatcaatctgtgttcgatgcacctcttcattataattttcattatgcttcaacagg
gaatggaaattatgatatgaggaatattttaaagggtacggtagtagaaagtcatcctacacttgctgttactcttgttgagaatcatgattctcagcc
tggacagtcattagaatctgttgtgagtccttggtttaagccgttggcctatgcatttatttaacgcgtgcagaagggtatccttctgttttttatggag
attactatggcacaaatggaaatagtagttatgaaattccaacgttaaaggataaaattgatccaattctgacggcacgaaaaaactttgcatatgg
tacgcaacatgattatttagaccatccagatgtgattggctggacaagagaaggggatagtatacatgctaattctggtttagcaacattaatctctg
atggaccaggaggatcaaaatggatgaatgttggaaagaacaacgcaggggaaatatggtacgatattacgggcaatcaaacaaatactgtaa
cgattaataaagatggatgggggcagttccatgtaaatgggggctctgtttcaatatatgttcagaagtaa

SEQ ID NO: 166

Met Gln Tyr Phe Glu Trp Tyr Val Pro Asn Asp Gly Glu His Trp Asn Arg Leu Arg Asn Asp Ala
Glu Asn Leu Ala His Lys Gly Ile Thr Ser Val Trp Ile Pro Pro Val Tyr Lys Gly Thr Ser Gln Asn
Asp Val Gly Tyr Gly Val Tyr Asp Val Tyr Asp Leu Gly Glu Phe Asn Gln Lys Gly Thr Ile Arg Thr
Lys Tyr Gly Thr Lys Ala Gln Leu Lys Ser Ala Ile Glu Ala Leu His Asn Gln Asn Ile Asp Val Tyr
Gly Asp Val Val Met Asn His Lys Gly Gly Ala Asp Tyr Thr Glu Val Val Thr Ala Val Glu Val Asp
Arg Asn Asn Arg Asn Ile Glu Thr Ser Ser Asp Tyr Gln Ile Asp Ala Trp Thr Gly Phe Asp Phe Pro
Gly Arg Arg Asp Ser Tyr Ser Asn Phe Lys Trp Arg Trp Phe His Phe Asp Gly Thr Asp Trp Asp Glu
Gly Arg Lys Leu Asn Arg Ile Tyr Lys Phe Lys Gly Val Gly Lys Ala Trp Asp Trp Glu Val Ser Ser
Glu Asn Gly Asn Tyr Asp Tyr Leu Met Tyr Ala Asp Leu Asp Phe Asp His Pro Glu Val Ala Asn
Glu Met Lys Asn Trp Gly Thr Trp Tyr Ala Asp Glu Leu Asn Leu Asp Gly Phe Arg Leu Asp Ala
Val Lys His Ile Asp His Glu Tyr Leu Arg Asp Trp Val Asn His Val Arg Lys Gln Thr Gly Lys Glu
Met Phe Thr Val Ala Glu Tyr Trp Gln Asn Asp Ile Arg Thr Leu Asn Asn Tyr Leu Gly Lys Val Asn
Tyr Asn Gln Ser Val Phe Asp Ala Pro Leu His Tyr Asn Phe His Tyr Ala Ser Thr Gly Asn Gly Asn
Tyr Asp Met Arg Asn Ile Leu Lys Gly Thr Val Val Glu Ser His Pro Thr Leu Ala Val Thr Leu Val
Glu Asn His Asp Ser Gln Pro Gly Gln Ser Leu Glu Ser Val Val Ser Pro Trp Phe Lys Pro Leu Ala
Tyr Ala Phe Ile Leu Thr Arg Ala Glu Gly Tyr Pro Ser Val Phe Tyr Gly Asp Tyr Tyr Gly Thr Asn
Gly Asn Ser Ser Tyr Glu Ile Pro Thr Leu Lys Asp Lys Ile Asp Pro Ile Leu Thr Ala Arg Lys Asn
Phe Ala Tyr Gly Thr Gln His Asp Tyr Leu Asp His Pro Asp Val Ile Gly Trp Thr Arg Glu Gly Asp
Ser Ile His Ala Asn Ser Gly Leu Ala Thr Leu Ile Ser Asp Gly Pro Gly Gly Ser Lys Trp Met Asn
Val Gly Lys Asn Asn Ala Gly Glu Ile Trp Tyr Asp Ile Thr Gly Asn Gln Thr Asn Thr Val Thr Ile
Asn Lys Asp Gly Trp Gly Gln Phe His Val Asn Gly Gly Ser Val Ser Ile Tyr Val Gln Lys

SEQ ID NO: 167

atgcaaacgattgcaaaaaaaggggatgaaacgatgaaagggaaaaaatggacagctttagctctaacactgccgctggctgctagcttatca
acaggcgttcacgccgaaaccgtacataaaggtaaatctccaacagcagataaaaacggtgtattttatgaggtgtatgtaaactctttttacgatg
caaataaagatggacatggtgatttaaaaggtcttacacaaaagttggattatttaaatgatggcaattctcatacaaagaatgatcttcaagtaaac
gggatttggatgatgccggtcaacccttctcccagctatcataaatatgatgtaacggactattataatattgatccgcagtatggaaatctgcaag
attttcgcaaactgatgaaagaagcagataaacgagatgtaaaagtcattatggacctcgttgtgaatcatacgagcagtgaacacccttggtttc
aagctgcattaaaagataaaaacagcaagtacagagattactatatctgggctgataaaaataccgacttgaatgaaaaaggatcttggggaca
gcaagtatggcataaagccccaaacggagagtatttttacggaacgttttgggaaggaatgccggacttaaattacgataatcctgaagtaagaa
aagaaatgattaacgtaggaaagttttggctaaagcaaggagttgacgggttccgtctagatgctgcgcttcatattttaaaggccaaacacctg

FIGURE 16SSS

aaggcgctaagaaaaatctcctgtggtggaatgaatttagagatgcaatgaaaaaggaaaaccctaacgtatatctaacgggtgaagtatggga
tcaaccggaagtagtagctccttactatcaatcgcttgattctttatttaactttgatttagcaggaaagattgtaaactctgtaaaatcaggaaatgat
caaggaatcgcgactgcagcagcggcaacggatgaactgttcaaatcatacaatccaaataaaattgacggtattttcttaaccaaccatgaccа
aaatcgcgtcatgagtgagctaagcggcgatgtgaataaagcaaagtcagctgcctctatcttacttacgcttcctggcaacccgtatatttattac
ggtgaagaaatcggcatgaccggtgaaaagcctgatgagttaatccgtgaaccgttccgctggtacgaaggaaacggacttggacaaaccag
ctgggaaacacctgtatacaacaaaggcggcaacggcgtgtctgtagaagcacaaacaaaacaaaaggactctttgttaaatcattaccgtgaa
atgattcgcgtgcgtcagcagcacgaagagttagtaaaaggaacgcttcaatctatttcagtagacagtaaagaagtcgttgcctatagccgcac
gtataaaggcaaatcgattagcgtgtatcataatatttcaaatcaaccggtaaaagtatctgtagcagcaaaaggtaaattgatttttggtagtgaaa
aaggtgctaagaaagtcaaaaatcagcttgtgattccggcgaatacaacggttttaataaaataa

SEQ ID NO: 168
Met Gln Thr Ile Ala Lys Lys Gly Asp Glu Thr Met Lys Gly Lys Lys Trp Thr Ala Leu Ala Leu Thr
Leu Pro Leu Ala Ala Ser Leu Ser Thr Gly Val His Ala Glu Thr Val His Lys Gly Lys Ser Pro Thr
Ala Asp Lys Asn Gly Val Phe Tyr Glu Val Tyr Val Asn Ser Phe Tyr Asp Ala Asn Lys Asp Gly His
Gly Asp Leu Lys Gly Leu Thr Gln Lys Leu Asp Tyr Leu Asn Asp Gly Asn Ser His Thr Lys Asn
Asp Leu Gln Val Asn Gly Ile Trp Met Met Pro Val Asn Pro Ser Pro Ser Tyr His Lys Tyr Asp Val
Thr Asp Tyr Tyr Asn Ile Asp Pro Gln Tyr Gly Asn Leu Gln Asp Phe Arg Lys Leu Met Lys Glu Ala
Asp Lys Arg Asp Val Lys Val Ile Met Asp Leu Val Val Asn His Thr Ser Ser Glu His Pro Trp Phe
Gln Ala Ala Leu Lys Asp Lys Asn Ser Lys Tyr Arg Asp Tyr Tyr Ile Trp Ala Asp Lys Asn Thr Asp
Leu Asn Glu Lys Gly Ser Trp Gly Gln Gln Val Trp His Lys Ala Pro Asn Gly Glu Tyr Phe Tyr Gly
Thr Phe Trp Glu Gly Met Pro Asp Leu Asn Tyr Asp Asn Pro Glu Val Arg Lys Glu Met Ile Asn Val
Gly Lys Phe Trp Leu Lys Gln Gly Val Asp Gly Phe Arg Leu Asp Ala Ala Leu His Ile Phe Lys Gly
Gln Thr Pro Glu Gly Ala Lys Lys Asn Leu Leu Trp Trp Asn Glu Phe Arg Asp Ala Met Lys Lys
Glu Asn Pro Asn Val Tyr Leu Thr Gly Glu Val Trp Asp Gln Pro Glu Val Val Ala Pro Tyr Tyr Gln
Ser Leu Asp Ser Leu Phe Asn Phe Asp Leu Ala Gly Lys Ile Val Asn Ser Val Lys Ser Gly Asn Asp
Gln Gly Ile Ala Thr Ala Ala Ala Ala Thr Asp Glu Leu Phe Lys Ser Tyr Asn Pro Asn Lys Ile Asp
Gly Ile Phe Leu Thr Asn His Asp Gln Asn Arg Val Met Ser Glu Leu Ser Gly Asp Val Asn Lys Ala
Lys Ser Ala Ala Ser Ile Leu Leu Thr Leu Pro Gly Asn Pro Tyr Ile Tyr Tyr Gly Glu Glu Ile Gly Met
Thr Gly Glu Lys Pro Asp Glu Leu Ile Arg Glu Pro Phe Arg Trp Tyr Glu Gly Asn Gly Leu Gly Gln
Thr Ser Trp Glu Thr Pro Val Tyr Asn Lys Gly Gly Asn Gly Val Ser Val Glu Ala Gln Thr Lys Gln
Lys Asp Ser Leu Leu Asn His Tyr Arg Glu Met Ile Arg Val Arg Gln Gln His Glu Glu Leu Val Lys
Gly Thr Leu Gln Ser Ile Ser Val Asp Ser Lys Glu Val Val Ala Tyr Ser Arg Thr Tyr Lys Gly Lys
Ser Ile Ser Val Tyr His Asn Ile Ser Asn Gln Pro Val Lys Val Ser Val Ala Ala Lys Gly Lys Leu Ile
Phe Gly Ser Glu Lys Gly Ala Lys Lys Val Lys Asn Gln Leu Val Ile Pro Ala Asn Thr Thr Val Leu
Ile Lys

SEQ ID NO: 169
atgaaaacattaaattaaaacgcacttttacgctaaccttgctgctcagtgctcctgccttgctgggcaaaatggcaccatgatgcagtatttt
cattggtatgtacctaatgatggcgcattatggacgcaggttgaaagcaatgctccagcactcgctgaaaacggttttacagcgctctggctacc
gccagcttacaaaggcgcgggcggcagtaatgacgtcggttatggcgtctatgatatgtacgatttaggtgagtttgatcaaaaaggctcagtac
gaaccaaatacggcaccaaggctcagtacatctctgcaatcaatgccgcgcacaacaacaatatccaaatctacggcgatgttgtgtttaaccac
cgaggtggtgctgatgggaagtcgtgggtcgataccaagcgcgttgattgggacaaccgtaacattgaactgggcgacaaatggattgaagct
tgggttgagtttaattttcctggccgcaacgacaaatactcaaacttccattggacttggtatcactttgacggtgttgactgggatgatgccggcaa
agaaaaagcgatctttaaattcaaaggcgaaggaaaagcatgggattgggaagtcagctctgaaaaaggcaattacgactacctaatgtacgc
cgatttagacatggatcaccaagaagttaaacaagagctgaaagattggggtgagtggtacatcaacatgaccggcgttgatggctttagaatg
gatgccgtgaagcacattaaatatcagtatctacaagagtggattgatcatttacgttggaaaacaggcaaagagcttttcaccgttggtgagtatt
ggaattacgacgtaaatcaactgcataactttattactaagacctctggcagtatgtcgttgttcgatgcgccgcttcacatgaacttctacaacgcg
tcaaaatctggcggcaattacgatatgcgccaaatcatgaatggcacgttgatgaaggacaacccagtcaaagctgtgactctcgtagaaaacc
acgatacacagccattgcaggcgttagagtcgacagtggattggtggttcaagcctcttgcttacgcattcattttattgcgtgaagaaggttatcc
atcagtgttctacgcagattactacggcgcgcagtacagcgacaaaggctacaacatcaatatggccaaagttccttacattgaagaacttgtaa

FIGURE 16TTT

cactgcgtaaagagtatgcgtatggcaaacagaattcttatctcgaccactgggatgtgattggctggacccgagagggcgatgctgaacatcc
aaactcaatggcggtgatcatgagtgatggaccaggtggcaaaaaatggatgtataccggtaagccaagcacgcgctatgtcgacaagctgg
gtatccgaactgaagaagtttggaccgataccaatggctgggcagaatttcctgtcaatggtggttcagtctcggtttgggtgggcgttaagtaa

SEQ ID NO: 170
Met Lys Thr Phe Lys Leu Lys Arg Thr Phe Leu Pro Leu Thr Leu Leu Leu Ser Ala Pro Ala Phe Ala
Gly Gln Asn Gly Thr Met Met Gln Tyr Phe His Trp Tyr Val Pro Asn Asp Gly Ala Leu Trp Thr Gln
Val Glu Ser Asn Ala Pro Ala Leu Ala Glu Asn Gly Phe Thr Ala Leu Trp Leu Pro Pro Ala Tyr Lys
Gly Ala Gly Gly Ser Asn Asp Val Gly Tyr Gly Val Tyr Asp Met Tyr Asp Leu Gly Glu Phe Asp
Gln Lys Gly Ser Val Arg Thr Lys Tyr Gly Thr Lys Ala Gln Tyr Ile Ser Ala Ile Asn Ala Ala His
Asn Asn Asn Ile Gln Ile Tyr Gly Asp Val Val Phe Asn His Arg Gly Gly Ala Asp Gly Lys Ser Trp
Val Asp Thr Lys Arg Val Asp Trp Asp Asn Arg Asn Ile Glu Leu Gly Asp Lys Trp Ile Glu Ala Trp
Val Glu Phe Asn Phe Pro Gly Arg Asn Asp Lys Tyr Ser Asn Phe His Trp Thr Trp Tyr His Phe Asp
Gly Val Asp Trp Asp Asp Ala Gly Lys Glu Lys Ala Ile Phe Lys Phe Lys Gly Glu Gly Lys Ala Trp
Asp Trp Glu Val Ser Ser Glu Lys Gly Asn Tyr Asp Tyr Leu Met Tyr Ala Asp Leu Asp Met Asp
His Gln Glu Val Lys Gln Glu Leu Lys Asp Trp Gly Glu Trp Tyr Ile Asn Met Thr Gly Val Asp Gly
Phe Arg Met Asp Ala Val Lys His Ile Lys Tyr Gln Tyr Leu Gln Glu Trp Ile Asp His Leu Arg Trp
Lys Thr Gly Lys Glu Leu Phe Thr Val Gly Glu Tyr Trp Asn Tyr Asp Val Asn Gln Leu His Asn
Phe Ile Thr Lys Thr Ser Gly Ser Met Ser Leu Phe Asp Ala Pro Leu His Met Asn Phe Tyr Asn Ala
Ser Lys Ser Gly Gly Asn Tyr Asp Met Arg Gln Ile Met Asn Gly Thr Leu Met Lys Asp Asn Pro Val
Lys Ala Val Thr Leu Val Glu Asn His Asp Thr Gln Pro Leu Gln Ala Leu Glu Ser Thr Val Asp Trp
Trp Phe Lys Pro Leu Ala Tyr Ala Phe Ile Leu Leu Arg Glu Glu Gly Tyr Pro Ser Val Phe Tyr Ala
Asp Tyr Tyr Gly Ala Gln Tyr Ser Asp Lys Gly Tyr Asn Ile Asn Met Ala Lys Val Pro Tyr Ile Glu
Glu Leu Val Thr Leu Arg Lys Glu Tyr Ala Tyr Gly Lys Gln Asn Ser Tyr Leu Asp His Trp Asp Val
Ile Gly Trp Thr Arg Glu Gly Asp Ala Glu His Pro Asn Ser Met Ala Val Ile Met Ser Asp Gly Pro
Gly Gly Lys Lys Trp Met Tyr Thr Gly Lys Pro Ser Thr Arg Tyr Val Asp Lys Leu Gly Ile Arg Thr
Glu Glu Val Trp Thr Asp Thr Asn Gly Trp Ala Glu Phe Pro Val Asn Gly Gly Ser Val Ser Val Trp
Val Gly Val Lys

SEQ ID NO: 171
gtgtatgtaaactcttttacgatgcaaataaagatggacatggtgatttaaaaggtcttacacaaaagttggattatttaaatgatggcaattctcata
caaagaatgatcttcaagtaaacgggatttggatgatgccggtcaacccttctcccagctatcataaatatgatgtaacggactattataatattgat
ccgcagtatggaaatctgcaagattttcgcaaactgatgaaagaagcagataaacgagatgtaaaagtcattatggacctcgttgtgaatcatac
gagcagtgaacacccttggtttcaagctgcattaaaagataaaaacagcaagtacagagattactatatctgggctgataaaaatacccgacttga
atgaaaaaggatcttggggacagcaagtatggcataaagccccaaacggagagtatttttacggaacgttttgggaaggaatgccggacttaaa
ttacgataatcctgaagtaagaaaagaaatgattaacgtaggaaagttttggctaaagcaaggagttgacgggttccgtctagatgctgcgcttca
tattttaaagggcaagcaacctgaaggcgctaagaaaaatctcctgtggtggaatgaatttagagatgcaatgaaaaaggaaaaccctaacgtat
atctaacgggtgaagtatgggatcaaccggaagtagtagctccttactatcaattcgttgattcttattaacttgatttagcaggaaagattgtaa
actctgtaaaatcaggaaatgatcaaggaatcgcgactgcagcagcggcaacggatgaactgttcaaatcatacaatccaaataaaattgacgg
tattttcttaaccaaccatgaccaaaatcgcgtcatgagtgagctaagcggcgatgtgaataaagcaaagtcagctgcctctatcttacttacgctt
cctggcaacccgtatatttattacggtgaagaaatcggcatgaccggtgaaaagcctgatgagttaatccgtgaaccgttccgctggtacgaagg
aaacggacttggacaaaccagctgggaaacacctgtatacaacaaaggcggcaacggcgtgtctgtagaagcacaaacaaaacaaaaggac
tctttgttaaatcattaccgtgaaatgattcgcgtgcgtcagcagcacgaagagttagtaaaaggaacgcttcaatctatttcagtagacagtaaag
aagtcgttgcctatagccgcacgtataaaggcaaatcgattagcgtgtatcataatatttcaaatcaaccggtaaaagtatctgtagcagcaaaag
gtaaattgatttttggtagtgaaaaaggtgctaagaaagtcaaaaatcagcttgtgattccggcgaatacaacggttttaataaaataa

SEQ ID NO: 172
Val Tyr Val Asn Ser Phe Tyr Asp Ala Asn Lys Asp Gly His Gly Asp Leu Lys Gly Leu Thr Gln
Lys Leu Asp Tyr Leu Asn Asp Gly Asn Ser His Thr Lys Asn Asp Leu Gln Val Asn Gly Ile Trp
Met Met Pro Val Asn Pro Ser Pro Ser Tyr His Lys Tyr Asp Val Thr Asp Tyr Tyr Asn Ile Asp Pro

FIGURE 16UUU

Gln Tyr Gly Asn Leu Gln Asp Phe Arg Lys Leu Met Lys Glu Ala Asp Lys Arg Asp Val Lys Val Ile Met Asp Leu Val Val Asn His Thr Ser Ser Glu His Pro Trp Phe Gln Ala Ala Leu Lys Asp Lys Asn Ser Lys Tyr Arg Asp Tyr Tyr Ile Trp Ala Asp Lys Asn Thr Asp Leu Asn Glu Lys Gly Ser Trp Gly Gln Gln Val Trp His Lys Ala Pro Asn Gly Glu Tyr Phe Tyr Gly Thr Phe Trp Glu Gly Met Pro Asp Leu Asn Tyr Asp Asn Pro Glu Val Arg Lys Glu Met Ile Asn Val Gly Lys Phe Trp Leu Lys Gln Gly Val Asp Gly Phe Arg Leu Asp Ala Ala Leu His Ile Phe Lys Gly Gln Thr Pro Glu Gly Ala Lys Lys Asn Leu Leu Trp Trp Asn Glu Phe Arg Asp Ala Met Lys Lys Glu Asn Pro Asn Val Tyr Leu Thr Gly Glu Val Trp Asp Gln Pro Glu Val Val Ala Pro Tyr Tyr Gln Ser Leu Asp Ser Leu Phe Asn Phe Asp Leu Ala Gly Lys Ile Val Asn Ser Val Lys Ser Gly Asn Asp Gln Gly Ile Ala Thr Ala Ala Ala Ala Thr Asp Glu Leu Phe Lys Ser Tyr Asn Pro Asn Lys Ile Asp Gly Ile Phe Leu Thr Asn His Asp Gln Asn Arg Val Met Ser Glu Leu Ser Gly Asp Val Asn Lys Ala Lys Ser Ala Ala Ser Ile Leu Leu Thr Leu Pro Gly Asn Pro Tyr Ile Tyr Tyr Gly Glu Glu Ile Gly Met Thr Gly Glu Lys Pro Asp Glu Leu Ile Arg Glu Pro Phe Arg Trp Tyr Glu Gly Asn Gly Leu Gly Gln Thr Ser Trp Glu Thr Pro Val Tyr Asn Lys Gly Gly Asn Gly Val Ser Val Glu Ala Gln Thr Lys Gln Lys Asp Ser Leu Leu Asn His Tyr Arg Glu Met Ile Arg Val Arg Gln Gln His Glu Glu Leu Val Lys Gly Thr Leu Gln Ser Ile Ser Val Asp Ser Lys Glu Val Val Ala Tyr Ser Arg Thr Tyr Lys Gly Lys Ser Ile Ser Val Tyr His Asn Ile Ser Asn Gln Pro Val Lys Val Ser Val Ala Ala Lys Gly Lys Leu Ile Phe Gly Ser Glu Lys Gly Ala Lys Lys Val Lys Asn Gln Leu Val Ile Pro Ala Asn Thr Thr Val Leu Ile Lys

SEQ ID NO: 173

atgcaaacgattgcaaaaaaaggggatgaaacgatgaaagggaaaaaatggacagctttagctctaacactgccgctggctgctagcttatca
acaggcgttcacgcagaaactgtacataaaggtaaagctccaacagcagataaaaacggtgtttttatgaggtgtatgtaaactcttttacgatg
caaataaagatggacatggtgatttaaaaggtctgacacaaaagttggattatttaaatgacggcaattctcatacaaagaatgatcttcaagtaaa
cgggatttggatgatgccggtaaaccccttctcctagctatcataaatatgatgtaacggactattataacattgatcctcagtacggaagtctgcaa
gatttccgcaaactgatgaaagaagcagataaacgagacgtaaaagttattatggaccttgttgtgaatcatacgagcagtgaacacccttggttt
caagctgcactaaaagataaaaacagcaagtacagagattactatatttgggctgataaaaataccgatttgaatgaaaaaggatcttggggaca
gcaagtatggcataaagctccaaacggagagtatttttacggaacgttctgggaaggaatgcctgacttaaattacgataaccctgaagtaagaa
aagaaatgattaacgtcggaaagttttggctaaagcaaggcgttgatggcttccgcttagatgctgcccttcatatctttaaaggtcaaactcctga
aggcgctaagaaaaatctcctgtggtggaatgagtttagagatgcaatgaaaaaagaaaaccctaacgtatatctaacgggtgaagtatgggat
cagccggaagtagtagctccttattatcaatcgcttgattccctatttaactttgatttagcaggaaaaattgtcagctctgtaaaagcaggaaatgat
caaggaatcgccactgcagcagcggcaacggatgagctgttcaaatcatacaatccaaataaaattgacggcattttcttaaccaaccatgacca
aaaccgcgtcatgagtgagctaagcggagatgtgaataaagcaaaatcagctgcttctatcttacttacgcttcctggaaatccgtatatttattacg
gtgaagaaattggcatgaccggtgaaaagcctgatgaattaatccgtgaaccgttccgctggtacgaaggcaacggaattggacaaactagct
gggaaacacctgtatataacaaaggcggcaatggtgtgtctgtagaagcacaaaccaaacaaaaggattcttgttaaatcattaccgtgaaatg
attcgcgtgcgtcagcagcacgaagagttagtaaaaggaacgcttcagtctatttcagtagacagtaaagaagttgtcgcttatagccgtacgtat
aaaggcaactccattagtgtgtatcataatatttcaaatcaacctgtaaaagtatctgtagcggcgaaaggtaaattgatttttgctagtgaaaaagg
tgctaaaaaaggcaaaaatcagcttgtgattcggcgaatgcgacggttttaataaataa

SEQ ID NO: 174

Met Gln Thr Ile Ala Lys Lys Gly Asp Glu Thr Met Lys Gly Lys Lys Trp Thr Ala Leu Ala Leu Thr Leu Pro Leu Ala Ala Ser Leu Ser Thr Gly Val His Ala Glu Thr Val His Lys Gly Lys Ala Pro Thr Ala Asp Lys Asn Gly Val Phe Tyr Glu Val Tyr Val Asn Ser Phe Tyr Asp Ala Asn Lys Asp Gly His Gly Asp Leu Lys Gly Leu Thr Gln Lys Leu Asp Tyr Leu Asn Asp Gly Asn Ser His Thr Lys Asn Asp Leu Gln Val Asn Gly Ile Trp Met Met Pro Val Asn Pro Ser Pro Ser Tyr His Lys Tyr Asp Val Thr Asp Tyr Tyr Asn Ile Asp Pro Gln Tyr Gly Ser Leu Gln Asp Phe Arg Lys Leu Met Lys Glu Ala Asp Lys Arg Asp Val Lys Val Ile Met Asp Leu Val Val Asn His Thr Ser Ser Glu His Pro Trp Phe Gln Ala Ala Leu Lys Asp Lys Asn Ser Lys Tyr Arg Asp Tyr Tyr Ile Trp Ala Asp Lys Asn Thr Asp Leu Asn Glu Lys Gly Ser Trp Gly Gln Gln Val Trp His Lys Ala Pro Asn Gly Glu Tyr Phe Tyr Gly Thr Phe Trp Glu Gly Met Pro Asp Leu Asn Tyr Asp Asn Pro Glu Val Arg Lys Glu Met Ile Asn Val Gly Lys Phe Trp Leu Lys Gln Gly Val Asp Gly Phe Arg Leu Asp Ala Ala Leu His Ile Phe Lys Gly

FIGURE 16VVV

Gln Thr Pro Glu Gly Ala Lys Lys Asn Leu Leu Trp Trp Asn Glu Phe Arg Asp Ala Met Lys Lys Glu Asn Pro Asn Val Tyr Leu Thr Gly Glu Val Trp Asp Gln Pro Glu Val Val Ala Pro Tyr Tyr Gln Ser Leu Asp Ser Leu Phe Asn Phe Asp Leu Ala Gly Lys Ile Val Ser Ser Val Lys Ala Gly Asn Asp Gln Gly Ile Ala Thr Ala Ala Ala Ala Thr Asp Glu Leu Phe Lys Ser Tyr Asn Pro Asn Lys Ile Asp Gly Ile Phe Leu Thr Asn His Asp Gln Asn Arg Val Met Ser Glu Leu Ser Gly Asp Val Asn Lys Ala Lys Ser Ala Ala Ser Ile Leu Leu Thr Leu Pro Gly Asn Pro Tyr Ile Tyr Tyr Gly Glu Glu Ile Gly Met Thr Gly Glu Lys Pro Asp Glu Leu Ile Arg Glu Pro Phe Arg Trp Tyr Glu Gly Asn Gly Ile Gly Gln Thr Ser Trp Glu Thr Pro Val Tyr Asn Lys Gly Gly Asn Gly Val Ser Val Glu Ala Gln Thr Lys Gln Lys Asp Ser Leu Leu Asn His Tyr Arg Glu Met Ile Arg Val Arg Gln Gln His Glu Glu Leu Val Lys Gly Thr Leu Gln Ser Ile Ser Val Asp Ser Lys Glu Val Val Ala Tyr Ser Arg Thr Tyr Lys Gly Asn Ser Ile Ser Val Tyr His Asn Ile Ser Asn Gln Pro Val Lys Val Ser Val Ala Ala Lys Gly Lys Leu Ile Phe Ala Ser Glu Lys Gly Ala Lys Lys Gly Lys Asn Gln Leu Val Ile Pro Ala Asn Ala Thr Val Leu Ile Lys

SEQ ID NO: 175

atgaaaaatataatacgactttgtgctgccagcgctatcctcacggtgtcccacgccagttacgccgacgcaattttacacgcgtttaactggcaat
ataccgatgtaaccgccaatgcaaatcaaattgccgcaaatggctttaaaaaagtcctcatttcacccgcaatgaaatccagcggcagtcaatgg
tgggcccgctatcaaccgcaagacttgcgtgtcattgattctccgctgggcaacaaacaagatttagtcgcgatgatcaatgcgctcaacagcgt
tggggtcgacgtgtatgctgacgtggtgcttaaccatatggctaacgagtcatggaagcgcagtgacctgaactacccggggagtgaggtgct
caacgactatcaatcccgcagtgcttactatcaaaggcaaacacttttcggcaatttacaggagaaccttttttccgagaatgatttccatccggca
ggctgtattaccaattggaatgatcctggccacgtccagtattggcgcttgtgcggcggacagggcgatactgggctaccggatctcgatcctaa
tcaatgggttgtgagtcagcagaagagttacttgaacgcactcaaatcaatgggaatcaaagggttccgtatcgatgcggtcaaacatatgagtc
aatatcaaatagaccaagtgtttaccccagacattaccgctggtatgcatatattcggagaagtcattaccagtggtgggcaaggtgatagcggct
atgaggcttttcttgcccttaccttaataataccgatcacgccgcttatgacttcccgctatttgcatcgattcgagccgcgttttcattctctggtgg
gttaaatcagctacacaatccacaagcctatggccaagcgttacaggactcacgtgcgatcacctttacgattacccacgacattccaaccaatg
acggtttccgctaccagatcatggatccaaccgatgaacagctcgcctatgcctacatcttgggcaaagatggaggaacgccacttgtctatagt
gatgacctacctgacagcgaagacaaagacagtggtcgttgggccgatgtgtggcaagatccgaacatgattaacatgcttgccttccacaacg
cgatgcaaggacaaagcatgactgtagtggctagcgatcaatgtaccttgctatttaagcgcggcaagcaaggcgtggtaggaatcaataaatg
tggcgagagtaagtcggtgactgtcgatacttaccagcatgagtttaactggtacaccccgtaccaagacgtattgagcggcgacatcaccaca
gtgagttctcgttatcaccaatttgttttgccagcgcgcagtgcaaggatgtggaaactataa

SEQ ID NO: 176

Met Lys Asn Ile Ile Arg Leu Cys Ala Ala Ser Ala Ile Leu Thr Val Ser His Ala Ser Tyr Ala Asp Ala Ile Leu His Ala Phe Asn Trp Gln Tyr Thr Asp Val Thr Ala Asn Ala Asn Gln Ile Ala Ala Asn Gly Phe Lys Lys Val Leu Ile Ser Pro Ala Met Lys Ser Ser Gly Ser Gln Trp Trp Ala Arg Tyr Gln Pro Gln Asp Leu Arg Val Ile Asp Ser Pro Leu Gly Asn Lys Gln Asp Leu Val Ala Met Ile Asn Ala Leu Asn Ser Val Gly Val Asp Val Tyr Ala Asp Val Val Leu Asn His Met Ala Asn Glu Ser Trp Lys Arg Ser Asp Leu Asn Tyr Pro Gly Ser Glu Val Leu Asn Asp Tyr Gln Ser Arg Ser Ala Tyr Tyr Gln Arg Gln Thr Leu Phe Gly Asn Leu Gln Glu Asn Leu Phe Ser Glu Asn Asp Phe His Pro Ala Gly Cys Ile Thr Asn Trp Asn Asp Pro Gly His Val Gln Tyr Trp Arg Leu Cys Gly Gly Gln Gly Asp Thr Gly Leu Pro Asp Leu Asp Pro Asn Gln Trp Val Val Ser Gln Gln Lys Ser Tyr Leu Asn Ala Leu Lys Ser Met Gly Ile Lys Gly Phe Arg Ile Asp Ala Val Lys His Met Ser Gln Tyr Gln Ile Asp Gln Val Phe Thr Pro Asp Ile Thr Ala Gly Met His Ile Phe Gly Glu Val Ile Thr Ser Gly Gly Gln Gly Asp Ser Gly Tyr Glu Ala Phe Leu Ala Pro Tyr Leu Asn Asn Thr Asp His Ala Ala Tyr Asp Phe Pro Leu Phe Ala Ser Ile Arg Ala Ala Phe Ser Phe Ser Gly Gly Leu Asn Gln Leu His Asn Pro Gln Ala Tyr Gly Gln Ala Leu Gln Asp Ser Arg Ala Ile Thr Phe Thr Ile Thr His Asp Ile Pro Thr Asn Asp Gly Phe Arg Tyr Gln Ile Met Asp Pro Thr Asp Glu Gln Leu Ala Tyr Ala Tyr Ile Leu Gly Lys Asp Gly Gly Thr Pro Leu Val Tyr Ser Asp Asp Leu Pro Asp Ser Glu Asp Lys Asp Ser Gly Arg Trp Ala Asp Val Trp Gln Asp Pro Asn Met Ile Asn Met Leu Ala Phe His Asn Ala Met Gln Gly Gln Ser Met Thr Val Val Ala Ser Asp Gln Cys Thr Leu Leu Phe Lys Arg Gly Lys Gln Gly Val Val Gly Ile Asn Lys Cys Gly

FIGURE 16WWW

Glu Ser Lys Ser Val Thr Val Asp Thr Tyr Gln His Glu Phe Asn Trp Tyr Thr Pro Tyr Gln Asp Val Leu Ser Gly Asp Ile Thr Thr Val Ser Ser Arg Tyr His Gln Phe Val Leu Pro Ala Arg Ser Ala Arg Met Trp Lys Leu

SEQ ID NO: 177

atgaaaacattcaaattaaaacgcacttttttaccgctgaccttgctgctcagtgctcctgcctttgctgggcaaaatggcaccatgatgcagtatttt
cattggtacgtacctaatgatggcgcattatggacgcaggttgaaagcaatgctccagtactcgctgaaaacggttttacagcgctctggctacc
gcccgcatacaaaggcgcgggcggcagtaatgacgtcggttatggcgtctatgatatgtacgatttaggtgagtttgaccaaaaaggctcagta
cgaaccaaatacggcaccaaggctcagtacatctctgcaatcaatgccgcgcacaacaacaatatccaaatttacggcgacgttgtgtttaacca
ccgaggtggcgctgatgggaagtcgtgggtcgataccaagcgcgttgattgggacaaccgcaatattgaactgggcgacaaatggattgaag
cttgggttgagtttaattttcctggccgcaacgacaaatactcgaacttccattggacttggtatcactttgacggtgttgactgggatgatgccggc
aaagaaaaagcgatctttaaattcaaaggcgaaggaaaagcatgggattgggaagtcagctctgaaaaaggcaattacgactacctaatgtac
gccgatttagacatggatcacccagaagttaaacaagagctgaaagattggggtgagtggtacatcaacatgaccggcgttgatggctttagaa
tggatgccgtgaagcacattaaatatcagtatctacaagagtggattgatcatttacgttggaaaacaggcaaagagcttttcaccgttggtgagta
ttggaattacgacgtaaatcaactgcacaactttattactaagacctctggcagtatgtcgttgttcgatgcgccgcttcacatgaatttctacaacgc
gtcaaaatctggcggcacttacgatatgcgccaaatcatgaatggcacgttgatgaaggacaacccagtcaaagcagtgactctcgtagaaaac
cacgatacgcagccattgcaggcgttagagtcgacagtagattggtggttcaagcctcttgcttacgcattcattttattgcgtgaagaaggttatc
catcggtgttctacgcagattactacggcgcgcagtacagcgacaaaggttacaacattaatatggccaaagtgccttacattgaagaacttgtaa
cactgcgtaaagagtatgcgtatggcaaacagaattcttatctcgaccattgggatgtgattggctggacccgagagggcgatgctgaacatcc
aaactcaatggcggtgatcatgagtgatggaccgggcggcacaaaatggatgtataccggtaagccaagtacgcgctatgtcgacaagctgg
gtatccgaactgaagatgtttggaccgatgccaatggctgggcagaatttcctgtcaatggtggttcagtctcggtttgggtgggcgttaagtaa

SEQ ID NO: 178

Met Lys Thr Phe Lys Leu Lys Arg Thr Phe Leu Pro Leu Thr Leu Leu Leu Ser Ala Pro Ala Phe Ala Gly Gln Asn Gly Thr Met Met Gln Tyr Phe His Trp Tyr Val Pro Asn Asp Gly Ala Leu Trp Thr Gln Val Glu Ser Asn Ala Pro Val Leu Ala Glu Asn Gly Phe Thr Ala Leu Trp Leu Pro Pro Ala Tyr Lys Gly Ala Gly Gly Ser Asn Asp Val Gly Tyr Gly Val Tyr Asp Met Tyr Asp Leu Gly Glu Phe Asp Gln Lys Gly Ser Val Arg Thr Lys Tyr Gly Thr Lys Ala Gln Tyr Ile Ser Ala Ile Asn Ala Ala His Asn Asn Asn Ile Gln Ile Tyr Gly Asp Val Val Phe Asn His Arg Gly Gly Ala Asp Gly Lys Ser Trp Val Asp Thr Lys Arg Val Asp Trp Asp Asn Arg Asn Ile Glu Leu Gly Asp Lys Trp Ile Glu Ala Trp Val Glu Phe Asn Phe Pro Gly Arg Asn Asp Lys Tyr Ser Asn Phe His Trp Thr Trp Tyr His Phe Asp Gly Val Asp Trp Asp Asp Ala Gly Lys Glu Lys Ala Ile Phe Lys Phe Lys Gly Glu Gly Lys Ala Trp Asp Trp Glu Val Ser Ser Glu Lys Gly Asn Tyr Asp Tyr Leu Met Tyr Ala Asp Leu Asp Met Asp His Pro Glu Val Lys Gln Glu Leu Lys Asp Trp Gly Glu Trp Tyr Ile Asn Met Thr Gly Val Asp Gly Phe Arg Met Asp Ala Val Lys His Ile Lys Tyr Gln Tyr Leu Gln Glu Trp Ile Asp His Leu Arg Trp Lys Thr Gly Lys Glu Leu Phe Thr Val Gly Glu Tyr Trp Asn Tyr Asp Val Asn Gln Leu His Asn Phe Ile Thr Lys Thr Ser Gly Ser Met Ser Leu Phe Asp Ala Pro Leu His Met Asn Phe Tyr Asn Ala Ser Lys Ser Gly Gly Thr Tyr Asp Met Arg Gln Ile Met Asn Gly Thr Leu Met Lys Asp Asn Pro Val Lys Ala Val Thr Leu Val Glu Asn His Asp Thr Gln Pro Leu Gln Ala Leu Glu Ser Thr Val Asp Trp Trp Phe Lys Pro Leu Ala Tyr Ala Phe Ile Leu Leu Arg Glu Glu Gly Tyr Pro Ser Val Phe Tyr Ala Asp Tyr Tyr Gly Ala Gln Tyr Ser Asp Lys Gly Tyr Asn Ile Asn Met Ala Lys Val Pro Tyr Ile Glu Glu Leu Val Thr Leu Arg Lys Glu Tyr Ala Tyr Gly Lys Gln Asn Ser Tyr Leu Asp His Trp Asp Val Ile Gly Trp Thr Arg Glu Gly Asp Ala Glu His Pro Asn Ser Met Ala Val Ile Met Ser Asp Gly Pro Gly Gly Thr Lys Trp Met Tyr Thr Gly Lys Pro Ser Thr Arg Tyr Val Asp Lys Leu Gly Ile Arg Thr Glu Asp Val Trp Thr Asp Ala Asn Gly Trp Ala Glu Phe Pro Val Asn Gly Gly Ser Val Ser Val Trp Val Gly Val Lys

SEQ ID NO: 179

atgaaaacattcaaattaaaacgcacttttttaccgctaaccttgctgctcagtgctcctgcctttgccgggcaaaatggcaccatgatgcagtactt
tcattggtacgtacctaatgatggcgcattatggacgcaggttgaaagcaatgctccagcactcgctgaaaacggttttacagcgctctggctacc

FIGURE 16XXX

gccagcttacaaaggcgcgggcggcagtaatgatgtcggttatggcgtctacgatatgtacgatttaggtgagtttgatcaaaaaggctcagtac
gaaccaaatacggtaccaaggctcagtacatctctgcaatcaatgctgcgcacaacaacaatatccaaatttacggcgacgttgtgtttaaccatc
gtggtggcgctgatgggaagtcgtgggtcgataccaagcgcgttgattgggacaaccgtaacattgaactgggcgacaaatggattgaagctt
gggttgagtttaattttcctagccgcaacgacaaatactcgaacttccattggacttggtatcactttgacggtgttgactgggatgatgccggcaa
agaaaaagcgatctttaaattcaaaggcgaaggaaaagcatgggattgggaagtcagctctgaaaaaggcaattacgactacctaatgtacgc
cgatttagacatggatcacccagaagttaaacaagagctgaaagattggggtgagtggtacatcaacatgaccggcgttgatggctttagaatg
gatgccgttaagcacattaaatatcagtatctacaagagtggattgatcatttacgttggaaaacaggcaaagagcttttcaccgttggtgagtattg
gaattacgacgtaaatcaactgcataactttattactaagacctctggcagtatgtcgttgttcgatgcgccgcttcacatgaacttctacaacgcgt
caaaatctggcggcaattacgatatgcgccaaatcatgaatggcacgttgatgaaggacaacccagtcaaagctgtgactctcgtagaaaacca
cgatacgcagccattgcaggcgttagagtcgacagtggattggtggttcaagcctcttgcttacgcattcatcttgttgcgtgaagaaggttatcca
tcggtgttctacgcagattactacggcgcgcagtacagcgacaaaggttacaacattaatatggccaaagtgccttacattgaagaacttgtaaca
ctgcgtaaagagtatgcgtatggcaaacagaattcttatctcgaccattgggatgtgattggctggactcgagagggcgatgctgaacatccaaa
ctcaatggcggtgatcatgagtgatggaccgggcggaacaaaatggatgtataccggtaatccaagcacgcgctatgtcgacaagctgggtat
ccgaactgaagatgtttggaccgatgccaatggctgggcagaatttcctgtcaatggtggttcagtctcggtttgggtgggcgttaagtaa

SEQ ID NO: 180

Met Lys Thr Phe Lys Leu Lys Arg Thr Phe Leu Pro Leu Thr Leu Leu Leu Ser Ala Pro Ala Phe Ala
Gly Gln Asn Gly Thr Met Met Gln Tyr Phe His Trp Tyr Val Pro Asn Asp Gly Ala Leu Trp Thr Gln
Val Glu Ser Asn Ala Pro Ala Leu Ala Glu Asn Gly Phe Thr Ala Leu Trp Leu Pro Pro Ala Tyr Lys
Gly Ala Gly Gly Ser Asn Asp Val Gly Tyr Gly Val Tyr Asp Met Tyr Asp Leu Gly Glu Phe Asp
Gln Lys Gly Ser Val Arg Thr Lys Tyr Gly Thr Lys Ala Gln Tyr Ile Ser Ala Ile Asn Ala Ala His
Asn Asn Asn Ile Gln Ile Tyr Gly Asp Val Val Phe Asn His Arg Gly Gly Ala Asp Gly Lys Ser Trp
Val Asp Thr Lys Arg Val Asp Trp Asp Asn Arg Asn Ile Glu Leu Gly Asp Lys Trp Ile Glu Ala Trp
Val Glu Phe Asn Phe Pro Ser Arg Asn Asp Lys Tyr Ser Asn Phe His Trp Thr Trp Tyr His Phe Asp
Gly Val Asp Trp Asp Asp Ala Gly Lys Glu Lys Ala Ile Phe Lys Phe Lys Gly Glu Gly Lys Ala Trp
Asp Trp Glu Val Ser Ser Glu Lys Gly Asn Tyr Asp Tyr Leu Met Tyr Ala Asp Leu Asp Met Asp
His Pro Glu Val Lys Gln Glu Leu Lys Asp Trp Gly Glu Trp Tyr Ile Asn Met Thr Gly Val Asp Gly
Phe Arg Met Asp Ala Val Lys His Ile Lys Tyr Gln Tyr Leu Gln Glu Trp Ile Asp His Leu Arg Trp
Lys Thr Gly Lys Glu Leu Phe Thr Val Gly Glu Tyr Trp Asn Tyr Asp Val Asn Gln Leu His Asn
Phe Ile Thr Lys Thr Ser Gly Ser Met Ser Leu Phe Asp Ala Pro Leu His Met Asn Phe Tyr Asn Ala
Ser Lys Ser Gly Gly Asn Tyr Asp Met Arg Gln Ile Met Asn Gly Thr Leu Met Lys Asp Asn Pro Val
Lys Ala Val Thr Leu Val Glu Asn His Asp Thr Gln Pro Leu Gln Ala Leu Glu Ser Thr Val Asp Trp
Trp Phe Lys Pro Leu Ala Tyr Ala Phe Ile Leu Leu Arg Glu Glu Gly Tyr Pro Ser Val Phe Tyr Ala
Asp Tyr Tyr Gly Ala Gln Tyr Ser Asp Lys Gly Tyr Asn Ile Asn Met Ala Lys Val Pro Tyr Ile Glu
Glu Leu Val Thr Leu Arg Lys Glu Tyr Ala Tyr Gly Lys Gln Asn Ser Tyr Leu Asp His Trp Asp Val
Ile Gly Trp Thr Arg Glu Gly Asp Ala Glu His Pro Asn Ser Met Ala Val Ile Met Ser Asp Gly Pro
Gly Gly Thr Lys Trp Met Tyr Thr Gly Asn Pro Ser Thr Arg Tyr Val Asp Lys Leu Gly Ile Arg Thr
Glu Asp Val Trp Thr Asp Ala Asn Gly Trp Ala Glu Phe Pro Val Asn Gly Gly Ser Val Ser Val Trp
Val Gly Val Lys

SEQ ID NO: 181

ttgccagaggccttcggcctggccattacgccgtcacatagccggcgggggaggttggtgggcgtgtcgcgcgggggcagcctgccgatgc
cggtcctccactggccggcgttcatcctcgtccggcgcttcgtcgccggtcatccgaacaagcacaagaaccggagtattgcgatgagccaca
ccctgcgtgccgccgtattggcggcgatcctgctgccgttcccgccctcgctgaccaggccggcaagagcccggccggcgtgcgctacca
cggcggcgacgaaatcatcctccagggcttccactggaacgtcgtccgcgaagcgcccaacgactggtacaacatccttcgccagcaggcct
cgacgatcgccgcggacggcttctcggcaatctggatgccggtgccctggcgtgacttctccagctggaccgacggcggcaagtcaggcgg
cggcgaaggctacttctggcacgacttcaacaagaacggccgctacggcagcgacgcccagctgcgccaggccgccggcgcactcggtgg
cgccggggtgaaggtgctctacgatgtggtgcccaatcacatgaaccgcggctatccggacaaggagatcaacctgccggccggccagggc
ttctggcgcaacgactgcaccgacccgggcaactaccccaacgactgcgatgacggtgaccgcttcatcggcggcaagtcggacctgaaca
ccggccatccgcagatctacggcatgtttcgcgacgagcttgccaacctgcgcagcgggtacggcgccggcggcttccgcttcgacttcgttc

FIGURE 16YYY

gcggctatgcgcccgaacgggtcgacagctggatgagcgacagcgccgacagcagtttctgcgttggcgagctgtggaaaagcccgtccga
gtacccgagctgggactggcgcaacacggcgagctggcagcagatcatcaaggactggtccgaccgggccaagtgcccggtgttcgacttc
gcgctcaaggagcgcatgcagaacggctcggtcgccgactggaagcatggcctcaatggcaacccggacccgcgctggcgcgaggtggc
ggtgacctttgtcgacaaccacgacaccggctattcgcccgggcagaacggcggccagcaccactgggcgctgcaggacgggctgatccg
ccaggcctacgcctacatcctcaccagcccgggcacgccggtggtgtactggtcgcacatgtacgactggggctacggcgacttcattcgcca
gctgatccaggtgcggcgcaccgctggcgtgcgcgccgattcggcgatcagcttccacagcggctacagcggcctggtcgctaccgtcagc
ggcagccatcagaccctggtggtggcgctcaactccgatctggccaaccccggccaggtcgccagcggcagcttcagcgaggcggtcaac
gccagcaacggccaggtgcgcgtctggcgcagcggtagcggcgatggcggcggcaatgacggcggcgagggcggtctggtcaatgtgaa
cttccgctgcgacaacggcgtgacgcagatgggcgacagcgtctacgcggtgggcaacgtcagccagctcggcaactggagcccggcctc
cgcggtacggctgaccgacaccagcagctatccgacctggaagggcagcatcgccctgcctgacggtcagaacgtggaatggaagtgcctg
atccgtaacgaggcggacgcgacgctggtgcgccagtggcaatcgggcggcaacaaccaggtccaggccgctgccggcgcgagcacca
gcggctcgttctga

SEQ ID NO: 182
Met Pro Glu Ala Phe Gly Leu Ala Ile Thr Pro Ser His Ser Arg Arg Gly Arg Leu Val Gly Val Ser
Arg Gly Gly Ser Leu Pro Met Pro Val Leu His Trp Pro Ala Phe Ile Leu Val Arg Arg Phe Val Ala
Gly His Pro Asn Lys His Lys Asn Arg Ser Ile Ala Met Ser His Thr Leu Arg Ala Ala Val Leu Ala
Ala Ile Leu Leu Pro Phe Pro Ala Leu Ala Asp Gln Ala Gly Lys Ser Pro Ala Gly Val Arg Tyr His
Gly Gly Asp Glu Ile Ile Leu Gln Gly Phe His Trp Asn Val Val Arg Glu Ala Pro Asn Asp Trp Tyr
Asn Ile Leu Arg Gln Gln Ala Ser Thr Ile Ala Ala Asp Gly Phe Ser Ala Ile Trp Met Pro Val Pro Trp
Arg Asp Phe Ser Ser Trp Thr Asp Gly Gly Lys Ser Gly Gly Gly Glu Gly Tyr Phe Trp His Asp Phe
Asn Lys Asn Gly Arg Tyr Gly Ser Asp Ala Gln Leu Arg Gln Ala Ala Gly Ala Leu Gly Gly Ala
Gly Val Lys Val Leu Tyr Asp Val Val Pro Asn His Met Asn Arg Gly Tyr Pro Asp Lys Glu Ile Asn
Leu Pro Ala Gly Gln Gly Phe Trp Arg Asn Asp Cys Thr Asp Pro Gly Asn Tyr Pro Asn Asp Cys
Asp Asp Gly Asp Arg Phe Ile Gly Gly Lys Ser Asp Leu Asn Thr Gly His Pro Gln Ile Tyr Gly Met
Phe Arg Asp Glu Leu Ala Asn Leu Arg Ser Gly Tyr Gly Ala Gly Gly Phe Arg Phe Asp Phe Val
Arg Gly Tyr Ala Pro Glu Arg Val Asp Ser Trp Met Ser Asp Ser Ala Asp Ser Ser Phe Cys Val Gly
Glu Leu Trp Lys Ser Pro Ser Glu Tyr Pro Ser Trp Asp Trp Arg Asn Thr Ala Ser Trp Gln Gln Ile Ile
Lys Asp Trp Ser Asp Arg Ala Lys Cys Pro Val Phe Asp Phe Ala Leu Lys Glu Arg Met Gln Asn
Gly Ser Val Ala Asp Trp Lys His Gly Leu Asn Gly Asn Pro Asp Pro Arg Trp Arg Glu Val Ala Val
Thr Phe Val Asp Asn His Asp Thr Gly Tyr Ser Pro Gly Gln Asn Gly Gly Gln His His Trp Ala Leu
Gln Asp Gly Leu Ile Arg Gln Ala Tyr Ala Tyr Ile Leu Thr Ser Pro Gly Thr Pro Val Val Tyr Trp Ser
His Met Tyr Asp Trp Gly Tyr Gly Asp Phe Ile Arg Gln Leu Ile Gln Val Arg Arg Thr Ala Gly Val
Arg Ala Asp Ser Ala Ile Ser Phe His Ser Gly Tyr Ser Gly Leu Val Ala Thr Val Ser Gly Ser His Gln
Thr Leu Val Val Ala Leu Asn Ser Asp Leu Ala Asn Pro Gly Gln Val Ala Ser Gly Ser Phe Ser Glu
Ala Val Asn Ala Ser Asn Gly Gln Val Arg Val Trp Arg Ser Gly Ser Gly Asp Gly Gly Gly Asn Asp
Gly Gly Glu Gly Gly Leu Val Asn Val Asn Phe Arg Cys Asp Asn Gly Val Thr Gln Met Gly Asp
Ser Val Tyr Ala Val Gly Asn Val Ser Gln Leu Gly Asn Trp Ser Pro Ala Ser Ala Val Arg Leu Thr
Asp Thr Ser Ser Tyr Pro Thr Trp Lys Gly Ser Ile Ala Leu Pro Asp Gly Gln Asn Val Glu Trp Lys
Cys Leu Ile Arg Asn Glu Ala Asp Ala Thr Leu Val Arg Gln Trp Gln Ser Gly Gly Asn Asn Gln Val
Gln Ala Ala Ala Gly Ala Ser Thr Ser Gly Ser Phe

SEQ ID NO: 183
atgcaaacgattgcaaaaaaaggggatgaaacgatgaaagggaaaaaatggacagctttagctctaacactgccgctggctgctagcttatca
acaggcgttcacgccgaaaccgtacataaaggtaagtctgaagcaacagataaaaacggtgtcttttatgaggtgtatgtaaactcttttacgata
caaataaagatggacatggtgatttaaaaggtctgacacaaaagttggattatttaaatgacggcaattctcatacaaagaatgatcttcaagtaaa
cgggatttggatgatgccagtcaaccccttctcctagctatcataaatatgatgtaacggactattataacattgatcctcagtacggaaatctgcaag
attttcgcaagctgatgaaagaagcagacaaacgagacgtaaaagtcattatggaccttgttgtgaatcatacgagcagcgaacaccctggttt
caagctgcattaaaagataaaaacagcaagtacagagattactatatttgggctgataaaaataccgatttgaatgaaaaaggatcttgggggca
gcaagtatggcataaagctccaaacggagagtatttttacggaacgttttgggaaggaatgcctgacttaaattacgataaccctgaagtaagaa

FIGURE 16ZZZ

aagaaatgattaacgtcggaaagttttggctaaagcaaggcgttaatggcttccgcttagatgctgcgcttcatatttttaaaggtcaaacacctga
aggcgctaagaaaaatatcctgtggtggaatgagtttagagatgcgatgaaaaaagaaaaccctaacgtatatctaacgggtgaagtatgggat
cagcctgaagtggtagctccttactatcaatcgcttgattctttatttaattttgatttagcaggaaaaattgtcagctctgtaaaagcaggaaatgatc
aaggaatcgccactgcagcagcggcaacagatgaactgttcaaatcatacaatccaaataaaattgacggcattttcttaaccaaccatgaccaa
aatcgcgtcatgagtgagctgagcggcgatgtgaacaaagcaaaatcagctgcttctatcttacttacgcttcctggcaacccgtatatttattacg
gtgaagaaattggcatgaccggtgaaaagcctgatgagttaatccgtgaaccattccgctggtacgaaggaaacggacttggacaaactagct
gggaaacacctgtatataacaaaggcggcaacggcgtgtctgtagaagtacaaaccaaacaaaaggattctttgttaaatcattatcgtgaaatg
attcgcgtgcgtcagcagcatgaagagttagtaaaaggaacgcttcaatctatttcagtagacagtaaagaagtggttgcctatagtcgcacgtat
aaaggcaactcgattagcgtgtatcataatatttcaaatcaacctgtaaaagtatctgtagcagcgaaaggtaaattgatttttgctagtgaaaaagg
tgctaaaaaagtcaaaaatcagcttgtaattccggctaatacaacggttttaataaaataa

SEQ ID NO: 184

Met Gln Thr Ile Ala Lys Lys Gly Asp Glu Thr Met Lys Gly Lys Lys Trp Thr Ala Leu Ala Leu Thr
Leu Pro Leu Ala Ala Ser Leu Ser Thr Gly Val His Ala Glu Thr Val His Lys Gly Lys Ser Glu Ala
Thr Asp Lys Asn Gly Val Phe Tyr Glu Val Tyr Val Asn Ser Phe Tyr Asp Thr Asn Lys Asp Gly His
Gly Asp Leu Lys Gly Leu Thr Gln Lys Leu Asp Tyr Leu Asn Asp Gly Asn Ser His Thr Lys Asn
Asp Leu Gln Val Asn Gly Ile Trp Met Met Pro Val Asn Pro Ser Pro Ser Tyr His Lys Tyr Asp Val
Thr Asp Tyr Tyr Asn Ile Asp Pro Gln Tyr Gly Asn Leu Gln Asp Phe Arg Lys Leu Met Lys Glu Ala
Asp Lys Arg Asp Val Lys Val Ile Met Asp Leu Val Val Asn His Thr Ser Ser Glu His Pro Trp Phe
Gln Ala Ala Leu Lys Asp Lys Asn Ser Lys Tyr Arg Asp Tyr Tyr Ile Trp Ala Asp Lys Asn Thr Asp
Leu Asn Glu Lys Gly Ser Trp Gly Gln Gln Val Trp His Lys Ala Pro Asn Gly Glu Tyr Phe Tyr Gly
Thr Phe Trp Glu Gly Met Pro Asp Leu Asn Tyr Asp Asn Pro Glu Val Arg Lys Glu Met Ile Asn Val
Gly Lys Phe Trp Leu Lys Gln Gly Val Asn Gly Phe Arg Leu Asp Ala Ala Leu His Ile Phe Lys Gly
Gln Thr Pro Glu Gly Ala Lys Lys Asn Ile Leu Trp Trp Asn Glu Phe Arg Asp Ala Met Lys Lys Glu
Asn Pro Asn Val Tyr Leu Thr Gly Glu Val Trp Asp Gln Pro Glu Val Val Ala Pro Tyr Tyr Gln Ser
Leu Asp Ser Leu Phe Asn Phe Asp Leu Ala Gly Lys Ile Val Ser Ser Val Lys Ala Gly Asn Asp Gln
Gly Ile Ala Thr Ala Ala Ala Ala Thr Asp Glu Leu Phe Lys Ser Tyr Asn Pro Asn Lys Ile Asp Gly
Ile Phe Leu Thr Asn His Asp Gln Asn Arg Val Met Ser Glu Leu Ser Gly Asp Val Asn Lys Ala Lys
Ser Ala Ala Ser Ile Leu Leu Thr Leu Pro Gly Asn Pro Tyr Ile Tyr Tyr Gly Glu Glu Ile Gly Met Thr
Gly Glu Lys Pro Asp Glu Leu Ile Arg Glu Pro Phe Arg Trp Tyr Glu Gly Asn Gly Leu Gly Gln Thr
Ser Trp Glu Thr Pro Val Tyr Asn Lys Gly Gly Asn Gly Val Ser Val Glu Val Gln Thr Lys Gln Lys
Asp Ser Leu Leu Asn His Tyr Arg Glu Met Ile Arg Val Arg Gln Gln His Glu Glu Leu Val Lys Gly
Thr Leu Gln Ser Ile Ser Val Asp Ser Lys Glu Val Val Ala Tyr Ser Arg Thr Tyr Lys Gly Asn Ser Ile
Ser Val Tyr His Asn Ile Ser Asn Gln Pro Val Lys Val Ser Val Ala Ala Lys Gly Lys Leu Ile Phe Ala
Ser Glu Lys Gly Ala Lys Lys Val Lys Asn Gln Leu Val Ile Pro Ala Asn Thr Thr Val Leu Ile Lys

SEQ ID NO: 185

atgaaactgatgaaagggaaaaaatggacagcttagctctaacaatgccgctggctgctagcttatcaacaggcgttcacgccgaaactgtac
ataaaggtaaagctccaacagcagataaaaacggtgtcttttatgaggtgtatgtaaactcttttacgatgcaaataaagatggacatggtgattta
aaaggtcttacacaaaagctggactatttaaatgacggaaattctcatacaaagaatgatcttcaagtaaacgggatttggatgatgccagtcaac
ccttctcctagctatcataaatatgatgtaacggattattataacattgatccgcagtacggaaatctgcaagattttcgcaagctgatgaaagaagc
agacaaacgagacgtaaaagtcattatggaccttgttgtgaatcatacgagcagcgaacaccccttggtttcaagctgcgttaaaagataaaaaca
gcaagtacagagattactatatttgggctgataaaaatacccgacttgaatgaaaaaggatcttggggacagcaagtatggcataaagctccaaac
ggagagtatttttacggaacgttttgggaaggaatgcctgacttaaattacgataaccctgaagtaagaaaagaaatgattaacgtcggaaagtttt
ggctaaagcaaggcgttgatggcttccgcttagatgctgcgcttcatatttttaaaggtcaaacgcctgaaggcgctaagaaaaatatctctgtggt
ggaatgagtttagagatgcgatgaaaaaagaaaaccctaacgtatatctaacgggtgaagtatgggatcagcctgaagtggtagctccttactat
caatcgcttgattccctatttaacttgatttagcagggaaaattgtcagttctgtaaaagcaggaaatgatcaaggaatcgccactgcagcagcgg
caacggatgagctgttcaaatcatacaatccaaataaaattgacggcattttcttaaccaaccatgaccaaaaccgcgtcatgagtgaactgatcg
gcgatgtgaacaaagcaaaatcagctgcttctatcttacttacgcttcctggcaacccgtatatttattacggtgaagaaattggcatgaccggtga
aaagcctgatgagttaatccgtgaaccgttccgctggtacgaaggaaacggacttggacaaaccagctgggaaacaccctgtatataacaaagg

FIGURE
16AAAA

cggcaacggcgtgtctgtagaagcacaaaccaaacaaaaggattctttgttaaatcattaccgtgaaatgattcgcgtgcgtcagcagcatgaag
agttagtaaaaggaacgcttcaatctatttagtagacagtaaagaagttgttgcctatagccgtacgtataaagacaactcgattagcgtgtatcat
aatatttcaaatcaaccggtaaaagtatctgtagcagcaaaaggtaaattaatttttgctagtgaaaaaggtgctaaaaaagtcaagaatcagcttg
tgattccggctaatacaacggttttaataaaataa

SEQ ID NO: 186
Met Lys Leu Met Lys Gly Lys Lys Trp Thr Ala Leu Ala Leu Thr Leu Pro Leu Ala Ala Ser Leu Ser
Thr Gly Val His Ala Glu Thr Val His Lys Gly Lys Ala Pro Thr Ala Asp Lys Asn Gly Val Phe Tyr
Glu Val Tyr Val Asn Ser Phe Tyr Asp Ala Asn Lys Asp Gly His Gly Asp Leu Lys Gly Leu Thr
Gln Lys Leu Asp Tyr Leu Asn Asp Gly Asn Ser His Thr Lys Asn Asp Leu Gln Val Asn Gly Ile Trp
Met Met Pro Val Asn Pro Ser Pro Ser Tyr His Lys Tyr Asp Val Thr Asp Tyr Tyr Asn Ile Asp Pro
Gln Tyr Gly Asn Leu Gln Asp Phe Arg Lys Leu Met Lys Glu Ala Asp Lys Arg Asp Val Lys Val
Ile Met Asp Leu Val Val Asn His Thr Ser Ser Glu His Pro Trp Phe Gln Ala Ala Leu Lys Asp Lys
Asn Ser Lys Tyr Arg Asp Tyr Tyr Ile Trp Ala Asp Lys Asn Thr Asp Leu Asn Glu Lys Gly Ser Trp
Gly Gln Gln Val Trp His Lys Ala Pro Asn Gly Glu Tyr Phe Tyr Gly Thr Phe Trp Glu Gly Met Pro
Asp Leu Asn Tyr Asp Asn Pro Glu Val Arg Lys Glu Met Ile Asn Val Gly Lys Phe Trp Leu Lys
Gln Gly Val Asp Gly Phe Arg Leu Asp Ala Ala Leu His Ile Phe Lys Gly Gln Thr Pro Glu Gly Ala
Lys Lys Asn Ile Leu Trp Trp Asn Glu Phe Arg Asp Ala Met Lys Lys Glu Asn Pro Asn Val Tyr Leu
Thr Gly Glu Val Trp Asp Gln Pro Glu Val Val Ala Pro Tyr Tyr Gln Ser Leu Asp Ser Leu Phe Asn
Phe Asp Leu Ala Gly Lys Ile Val Ser Ser Val Lys Ala Gly Asn Asp Gln Gly Ile Ala Thr Ala Ala
Ala Ala Thr Asp Glu Leu Phe Lys Ser Tyr Asn Pro Asn Lys Ile Asp Gly Ile Phe Leu Thr Asn His
Asp Gln Asn Arg Val Met Ser Glu Leu Ile Gly Asp Val Asn Lys Ala Lys Ser Ala Ala Ser Ile Leu
Leu Thr Leu Pro Gly Asn Pro Tyr Ile Tyr Tyr Gly Glu Glu Ile Gly Met Thr Gly Glu Lys Pro Asp
Glu Leu Ile Arg Glu Pro Phe Arg Trp Tyr Glu Gly Asn Gly Leu Gly Gln Thr Ser Trp Glu Thr Pro
Val Tyr Asn Lys Gly Gly Asn Gly Val Ser Val Glu Ala Gln Thr Lys Gln Lys Asp Ser Leu Leu Asn
His Tyr Arg Glu Met Ile Arg Val Arg Gln Gln His Glu Glu Leu Val Lys Gly Thr Leu Gln Ser Ile
Leu Val Asp Ser Lys Glu Val Val Ala Tyr Ser Arg Thr Tyr Lys Asp Asn Ser Ile Ser Val Tyr His
Asn Ile Ser Asn Gln Pro Val Lys Val Ser Val Ala Ala Lys Gly Lys Leu Ile Phe Ala Ser Glu Lys
Gly Ala Lys Lys Val Lys Asn Gln Leu Val Ile Pro Ala Asn Thr Thr Val Leu Ile Lys

SEQ ID NO: 187
ttgtatctcatccaggaggggcacatgcgttttccgcccattattcacccgcttaccggcctggccgttccggttggagctctgcgtaccgcacag
agctgcggcataggggagtttgccgacttgccggttcttgccgaattctgcaaaaaagccggatttgatcttgtacagcttcttccggtcaatgac
accggcacagaaagttctccatacagcgcgctttctgcctttgccctgcacccgctgtatatcaggctttccgacctgcctgaagcagcgggtttc
gaaaagcagattacagatctgaaaagccggtttgaggacttgcctcgtttcagctatacggagctgcgccgtgccaaactggatatcctgcgtgc
agtgtttgataaaaacaaggcaaccatcatcggcagtgccgaactggaagcctggatttcagataaccccctggatcatcgaatatgcggttttat
gaacccagaaacaccgcaacttttgaagccgggctggaaacattgggaaaagctgcgcaaccctacctctataaggaaatacaaaaaaccctggcag
ggtaaaaccttggaggctgaccatcaattcttgcatggctgcagatgcggctggaccagcagtttactgccgccgctacagagtgcaaogcc
ctgggtgtctatcttaagggcgatatacctataatgatgaacgaggattccgcagatgcctgggcgaatccggaattcttccgtgacgatcttcgg
gccggaagtcccctgacggtgaaaaccccagggacaaaactggggcttccccatttataactgggaaaaccttgcaaatgacgggtacag
ctggtggaaaaaacgtctgaagcacagcgcacggtattaccatgcctaccgcattgaccatattcttgggtttttccggatatgggctataccctat
ggcgaatactccggctacctgggatggcccttgccgcatgaaccggtaagtgcagcagaactggcagaacggggctttccaaggaccgctt
gcgctggcttaccgaaccccacttgcctacacgggcagccgaggaagcgaataactgggactatctgggaacacacggctatctgaatcaga
tcatgaaccgtatcggtgaagaagaactatggctgttcaagcccgagatcacctgcgaggcagatatacgaaacacaaacctgccggatgccc
tgaaagaggttctggtacggcagtggaaaaaccggctgctgcaggttaccggccgcgacgaaaaaggacggacaatctactatccgctgtgg
cgtttccgtgacagcactgcatggcagacgcttaccgatggcgagaaacactccctggaagagctgttcgcccaaaaagcggcgcacaatga
aaccctgtggcgagaacaggcggtggaacttctgggtgagctgacgcgatctacggatatgcttgcctgtgctgaagatctgggaagtattccc
cacagtgtaccggaagtgctttcaaaccttcaatttacagtctgcgggttacccgctgggcccgccaatgggatgcccccggccagccctttca
cagactggaggagtatccgctcatgtcggtagcgaccccatcggttcatgattcctctaccctgcgcggatggtgggaaaccgaaggcggcga
ccgggcctttatggacgcatggcctccggaacaggatgcatacgcaggagcaggccgccatgagttcgaaggcgcctgggggaccccgcca

Figure 16 (cont.)

FIGURE 16BBBB

ggcatcctgggtactccgtaaactctgcgaagcccgttccgcgctctgtgttttccccatccaggatattttggccctgtcttcagacttttatgcaat
gacagcggacgaggaacgcatcaatattccgggcagtgtatccggatttaactggacataccggttgcctgcggcaatcgaggatttatctaaa
aacagccaacttataaccgcaatccagaccgcgttgcaggaccgccgggcgaggaaggcacaaggagcacagcaatga

SEQ ID NO: 188

Met Tyr Leu Ile Gln Glu Gly His Met Arg Phe Pro Pro Ile Ile His Pro Leu Thr Gly Leu Ala Val Pro Val Gly Ala Leu Arg Thr Ala Gln Ser Cys Gly Ile Gly Glu Phe Ala Asp Leu Pro Val Leu Ala Glu Phe Cys Lys Lys Ala Gly Phe Asp Leu Val Gln Leu Leu Pro Val Asn Asp Thr Gly Thr Glu Ser Ser Pro Tyr Ser Ala Leu Ser Ala Phe Ala Leu His Pro Leu Tyr Ile Arg Leu Ser Asp Leu Pro Glu Ala Ala Gly Phe Glu Lys Gln Ile Thr Asp Leu Lys Ser Arg Phe Glu Asp Leu Pro Arg Phe Ser Tyr Thr Glu Leu Arg Arg Ala Lys Leu Asp Ile Leu Arg Ala Val Phe Asp Lys Asn Lys Ala Thr Ile Ile Gly Ser Ala Glu Leu Glu Ala Trp Ile Ser Asp Asn Pro Trp Ile Ile Glu Tyr Ala Val Phe Met Asn Gln Lys His Arg Asn Phe Glu Ala Gly Trp Lys His Trp Glu Lys Leu Arg Asn Pro Thr His Asn Glu Ile Gln Lys Thr Trp Gln Gly Lys Thr Trp Gln Ala Asp His Gln Phe Phe Ala Trp Leu Gln Met Arg Leu Asp Gln Gln Phe Thr Ala Ala Ala Thr Glu Cys Asn Ala Leu Gly Val Tyr Leu Lys Gly Asp Ile Pro Ile Met Met Asn Glu Asp Ser Ala Asp Ala Trp Ala Asn Pro Glu Phe Phe Arg Asp Asp Leu Arg Ala Gly Ser Pro Pro Asp Gly Glu Asn Pro Gln Gly Gln Asn Trp Gly Phe Pro Ile Tyr Asn Trp Glu Asn Leu Ala Asn Asp Gly Tyr Ser Trp Trp Lys Lys Arg Leu Lys His Ser Ala Arg Tyr Tyr His Ala Tyr Arg Ile Asp His Ile Leu Gly Phe Phe Arg Ile Trp Ala Ile Pro Tyr Gly Glu Tyr Ser Gly Tyr Leu Gly Trp Pro Leu Pro His Glu Pro Val Ser Ala Ala Glu Leu Ala Glu Arg Gly Phe Ser Lys Asp Arg Leu Arg Trp Leu Thr Glu Pro His Leu Pro Thr Arg Ala Ala Glu Glu Ala Asn Asn Trp Asp Tyr Leu Gly Thr His Gly Tyr Leu Asn Gln Ile Met Asn Arg Ile Gly Glu Glu Glu Leu Trp Leu Phe Lys Pro Glu Ile Thr Cys Glu Ala Asp Ile Arg Asn Thr Asn Leu Pro Asp Ala Leu Lys Glu Val Leu Val Arg Gln Trp Lys Asn Arg Leu Leu Gln Val Thr Gly Arg Asp Glu Lys Gly Arg Thr Ile Tyr Tyr Pro Leu Trp Arg Phe Arg Asp Ser Thr Ala Trp Gln Thr Leu Thr Asp Gly Glu Lys His Ser Leu Glu Glu Leu Phe Ala Gln Lys Ala Ala His Asn Glu Thr Leu Trp Arg Glu Gln Ala Val Glu Leu Leu Gly Glu Leu Thr Arg Ser Thr Asp Met Leu Ala Cys Ala Glu Asp Leu Gly Ser Ile Pro His Ser Val Pro Glu Val Leu Ser Asn Leu Ser Ile Tyr Ser Leu Arg Val Thr Arg Trp Ala Arg Gln Trp Asp Ala Pro Gly Gln Pro Phe His Arg Leu Glu Glu Tyr Pro Leu Met Ser Val Ala Thr Pro Ser Val His Asp Ser Ser Thr Leu Arg Gly Trp Trp Glu Thr Glu Gly Gly Asp Arg Ala Phe Met Asp Ala Trp Pro Pro Glu Gln Asp Ala Tyr Ala Gly Ala Gly Arg His Glu Phe Glu Gly Ala Trp Gly Pro Arg Gln Ala Ser Trp Val Leu Arg Lys Leu Cys Glu Ala Arg Ser Ala Leu Cys Val Phe Pro Ile Gln Asp Ile Leu Ala Leu Ser Ser Asp Phe Tyr Ala Met Thr Ala Asp Glu Glu Arg Ile Asn Ile Pro Gly Ser Val Ser Gly Phe Asn Trp Thr Tyr Arg Leu Pro Ala Ala Ile Glu Asp Leu Ser Lys Asn Ser Gln Leu Ile Thr Ala Ile Gln Thr Ala Leu Gln Asp Arg Arg Ala Arg Lys Ala Gln Gly Ala Gln Gln

SEQ ID NO: 189

atgcaaacgattgcaaaaaaaggggatgaaacgatgaaagggaaaaaatggacaagctttagctctaacactgccgctggctgctagcttatca
acaggcgttcacgccgaaaccgtacataaaggtaaatctccagctgcagataaaaacggtgtcttttatgaggtgtatgtaaactcttttacgatg
caaataaagatggacatggtgatttaaaaggtcttacacaaaaactggactatttaaatgatggcaattctcatacaaagaatgatcttcaagtaaa
cgggatttggatgatgccgatcaacccttctcctagctatcataaatatgatgtaacggactattataacattgattctcagtacggaaatctgcaag
attttcgcaagctaatgaaagaagcagataaacgagatgtaaaagttattatggacctcgttgtgaatcatacgagcagtgaacacccttggtttca
agctgcgttaaaagataaaaacagcaagtacagagattactatatttgggctgataaaaatacccgatttgaatgaaaaaggatcttggggacaac
aagtatggcacaaagctccaaacggagagtatttttacggaacgttctgggaaggaatgcctgacttaaattacgataaccctgaagtaagaaaa
gaaatgattaacgtcggaaagttttggctaaagcaaggcgttgacggcttccgcttagatgctgcccttcatatctttaaaggtcaaacacctgaa
ggcgctaagaaaaatattgtgtggtggaatgaatttagagatgcgatgaaaaaagaaaacccgaacgtatatctaacgggcgaagtatgggatc
agccggaagtggtagctccttattatcagtcgcttgattccctatttaactttgatttagcaggaaaaattgtcagctctgtaaaagcaggaaatgatc
aaggaatcgctactgcagcagcggcaacagatgaactgttcaaatcatacaatccaaataaaattgacggcattttcttaaccaatcatgaccaaa
atcgcgtcatgagtgagttaagcggagatgtcaataaagcaaagtcagctgcctctatcttacttacgcttcctggaaatccgtatatttattacggt
gaagaaatcggcatgaccggtgaaaagcctgatgaattaatccgtgaaccgttccgctggtacgaaggaaacggacttggacaaactagttgg

FIGURE 16CCCC

gaaacacctgtatacaataaaggcggcaacggcgtgtctgtagaagcacaaaccaaacaaaaggactctttgttaaatcattaccgtgaaatgat
tcgcgtgcgtcagcagcacgaagagttagtaaaaggaacgcttcaatctatttcagtagacagtaaagaagttgttgcttatagccgtacgtataa
aggcaactccattagtgtgtatcataatatttcaaatcaaccgtaaaagtatctgtagcagcgaaaggtaaattgatttttgctagtgaaaaaggtg
ctaaaaaggtcaaaaatcagcttgtgattccggcgaatacaacggttttagtaaaataa

SEQ ID NO: 190
Met Gln Thr Ile Ala Lys Lys Gly Asp Glu Thr Met Lys Gly Lys Lys Trp Thr Ala Leu Ala Leu Thr Leu Pro Leu Ala Ala Ser Leu Ser Thr Gly Val His Ala Glu Thr Val His Lys Gly Lys Ser Pro Ala Ala Asp Lys Asn Gly Val Phe Tyr Glu Val Tyr Val Asn Ser Phe Tyr Asp Ala Asn Lys Asp Gly His Gly Asp Leu Lys Gly Leu Thr Gln Lys Leu Asp Tyr Leu Asn Asp Gly Asn Ser His Thr Lys Asn Asp Leu Gln Val Asn Gly Ile Trp Met Met Pro Ile Asn Pro Ser Pro Ser Tyr His Lys Tyr Asp Val Thr Asp Tyr Tyr Asn Ile Asp Ser Gln Tyr Gly Asn Leu Gln Asp Phe Arg Lys Leu Met Lys Glu Ala Asp Lys Arg Asp Val Lys Val Ile Met Asp Leu Val Val Asn His Thr Ser Ser Glu His Pro Trp Phe Gln Ala Ala Leu Lys Asp Lys Asn Ser Lys Tyr Arg Asp Tyr Tyr Ile Trp Ala Asp Lys Asn Thr Asp Leu Asn Glu Lys Gly Ser Trp Gly Gln Gln Val Trp His Lys Ala Pro Asn Gly Glu Tyr Phe Tyr Gly Thr Phe Trp Glu Gly Met Pro Asp Leu Asn Tyr Asp Asn Pro Glu Val Arg Lys Glu Met Ile Asn Val Gly Lys Phe Trp Leu Lys Gln Gly Val Asp Gly Phe Arg Leu Asp Ala Ala Leu His Ile Phe Lys Gly Gln Thr Pro Glu Gly Ala Lys Lys Asn Ile Val Trp Trp Asn Glu Phe Arg Asp Ala Met Lys Lys Glu Asn Pro Asn Val Tyr Leu Thr Gly Glu Val Trp Asp Gln Pro Glu Val Val Ala Pro Tyr Tyr Gln Ser Leu Asp Ser Leu Phe Asn Phe Asp Leu Ala Gly Lys Ile Val Ser Ser Val Lys Ala Gly Asn Asp Gln Gly Ile Ala Thr Ala Ala Ala Ala Thr Asp Glu Leu Phe Lys Ser Tyr Asn Pro Asn Lys Ile Asp Gly Ile Phe Leu Thr Asn His Asp Gln Asn Arg Val Met Ser Glu Leu Ser Gly Asp Val Asn Lys Ala Lys Ser Ala Ala Ser Ile Leu Leu Thr Leu Pro Gly Asn Pro Tyr Ile Tyr Tyr Gly Glu Glu Ile Gly Met Thr Gly Glu Lys Pro Asp Glu Leu Ile Arg Glu Pro Phe Arg Trp Tyr Glu Gly Asn Gly Leu Gly Gln Thr Ser Trp Glu Thr Pro Val Tyr Asn Lys Gly Gly Asn Gly Val Ser Val Glu Ala Gln Thr Lys Gln Lys Asp Ser Leu Leu Asn His Tyr Arg Glu Met Ile Arg Val Arg Gln Gln His Glu Glu Leu Val Lys Gly Thr Leu Gln Ser Ile Ser Val Asp Ser Lys Glu Val Val Ala Tyr Ser Arg Thr Tyr Lys Gly Asn Ser Ile Ser Val Tyr His Asn Ile Ser Asn Gln Pro Val Lys Val Ser Val Ala Ala Lys Gly Lys Leu Ile Phe Ala Ser Glu Lys Gly Ala Lys Lys Val Lys Asn Gln Leu Val Ile Pro Ala Asn Thr Thr Val Leu Val Lys

SEQ ID NO: 191
atgcaaacgattgcaaaaaaaggggatgaaacgatgaaagggaaaaaatggacagctttagctctaacactgccgctggctgctagcttatca
acaggcgttcacgccgaaaccgtacataaaggtaaatctccaacagcagataaaaacggtgtctttatgaagtgtatgtaaactcttttacgatg
caaataaagatggacatggtgacttaaaaggtcttacacaaaagttggactatttaaatgacggcaattctcatacaaaaaatgatcttcaagtaaa
cgggatttggatgatgccagtcaacccttctcctagctatcataaatatgatgtaacggactattataacattgatccgcagtacggaaatctgcaa
gattttcgcaagctgatgaaagaagcagacaaacgagacgtaaaagtcattatggaccttgttgtgaatcatacgagcagtgaacaccccttggttt
caagctgcgttaaaagataaaaacaggcaagtaacaggattactatttggctgtaaaaatacccgacttgaatgaaaaaggatcttgggggtca
acaagtatggcataaagctccaaccggagagtatttttacggtacgttctgggaaggaatgccttgacttaaattacgataaccctgaagtaagaa
aagaaatgattaacgtcggaaagttttggctaaagcaaggcgttgacgggttccgcttagatgctgcgcttcatattttaaaggtcaaacagctga
aggcgctaagaaaaatatcctgtggtggaatgagtttagagatgcgatgaaaaaagaaaatccgaatgtatatctaacgggtgaagtatgggat
cagcctgaagtggtagctccttattatcaatcgcttgattctttatttaatttgatttagcaggaaaaattgtcagctctgtaaaagcaggaaatgatc
aaggaatcgccactgcagcagcagcaacagatgaactgttcaaatcatacaatccaaacaaaattgatggcatattcttaaccaaccatgaccaa
aatcgcgtcatgagtgagctgagcggcgatgtgagcaaagcaaaatcagctgcttctatcttacttacgcttcctggcaacccgtatatttattacg
gtgaagaaatcggcatgaccggtgaaaagcctgatgaattaatccgtgaaccgttccgctggtacgaaggaaacggacttggacaaaccagtt
gggaaacacctgtatacaataaaggcggaaacggtgtgtctgtagaagcacaaaccaaacaaaaggattctttgttaaatcattaccgtgaaatg
attcgcgtgcgtcagcagcatgaagagttagtaaaaggaacgcttcaatctatttcagtagacagtaaagaagttgttgcttatagccgtacgtata
aaggcaactccattagtgtgtatcataatatttcaaatcaaccggtaaaagtatctgtagcagcgaaaggtaaattgatttttgctagtgaaaaaggt
gctaagaaagtcaaaaatcagcttgtggttccggcgaatacaacggttttaatgaaataa

SEQ ID NO: 192

2013201807 18 Mar 2013

FIGURE 16DDDD

Met Gln Thr Ile Ala Lys Lys Gly Asp Glu Thr Met Lys Gly Lys Lys Trp Thr Ala Leu Ala Leu Thr Leu Pro Leu Ala Ala Ser Leu Ser Thr Gly Val His Ala Glu Thr Val His Lys Gly Lys Ser Pro Thr Ala Asp Lys Asn Gly Val Phe Tyr Glu Val Tyr Val Asn Ser Phe Tyr Asp Ala Asn Lys Asp Gly His Gly Asp Leu Lys Gly Leu Thr Gln Lys Leu Asp Tyr Leu Asn Asp Gly Asn Ser His Thr Lys Asn Asp Leu Gln Val Asn Gly Ile Trp Met Met Pro Val Asn Pro Ser Pro Ser Tyr His Lys Tyr Asp Val Thr Asp Tyr Tyr Asn Ile Asp Pro Gln Tyr Gly Asn Leu Gln Asp Phe Arg Lys Leu Met Lys Glu Ala Asp Lys Arg Asp Val Lys Val Ile Met Asp Leu Val Val Asn His Thr Ser Ser Glu His Pro Trp Phe Gln Ala Ala Leu Lys Asp Lys Asn Ser Lys Tyr Arg Asp Tyr Tyr Ile Trp Ala Asp Lys Asn Thr Asp Leu Asn Glu Lys Gly Ser Trp Gly Gln Gln Val Trp His Lys Ala Pro Asn Gly Glu Tyr Phe Tyr Gly Thr Phe Trp Glu Gly Met Pro Asp Leu Asn Tyr Asp Asn Pro Glu Val Arg Lys Glu Met Ile Asn Val Gly Lys Phe Trp Leu Lys Gln Gly Val Asp Gly Phe Arg Leu Asp Ala Ala Leu His Ile Phe Lys Gly Gln Thr Ala Glu Gly Ala Lys Lys Asn Ile Leu Trp Trp Asn Glu Phe Arg Asp Ala Met Lys Lys Glu Asn Pro Asn Val Tyr Leu Thr Gly Glu Val Trp Asp Gln Pro Glu Val Val Ala Pro Tyr Tyr Gln Ser Leu Asp Ser Leu Phe Asn Phe Asp Leu Ala Gly Lys Ile Val Ser Ser Val Lys Ala Gly Asn Asp Gln Gly Ile Ala Thr Ala Ala Ala Ala Thr Asp Glu Leu Phe Lys Ser Tyr Asn Pro Asn Lys Ile Asp Gly Ile Phe Leu Thr Asn His Asp Gln Asn Arg Val Met Ser Glu Leu Ser Gly Asp Val Ser Lys Ala Lys Ser Ala Ala Ser Ile Leu Leu Thr Leu Pro Gly Asn Pro Tyr Ile Tyr Tyr Gly Glu Glu Ile Gly Met Thr Gly Glu Lys Pro Asp Glu Leu Ile Arg Glu Pro Phe Arg Trp Tyr Glu Gly Asn Gly Leu Gly Gln Thr Ser Trp Glu Thr Pro Val Tyr Asn Lys Gly Gly Asn Gly Val Ser Val Glu Ala Gln Thr Lys Gln Lys Asp Ser Leu Leu Asn His Tyr Arg Glu Met Ile Arg Val Arg Gln Gln His Glu Glu Leu Val Lys Gly Thr Leu Gln Ser Ile Ser Val Asp Ser Lys Glu Val Val Ala Tyr Ser Arg Thr Tyr Lys Gly Asn Ser Ile Ser Val Tyr His Asn Ile Ser Asn Gln Pro Val Lys Val Ser Val Ala Ala Lys Gly Lys Leu Ile Phe Ala Ser Glu Lys Gly Ala Lys Lys Val Lys Asn Gln Leu Val Val Pro Ala Asn Thr Thr Val Leu Met Lys

SEQ ID NO: 193

atgaaattcaaaaagagtttatctgccgggctccttttgttcggaggtctgagcggtgtgacaccatccgtcgctgcggaggtgccacgaaccgc
atttgtccatttattcgaatggagttggccggatattgccaccgaatgcgaaaccttttcttggccctaaggggttctctgcggttcaggtgtctccgc
cgcaaaaaagcgtcagcaatgctgcctggtgggcgcgctaccaacctgttagttactcttttgaagggcgcagtggaacccgggctcaatttgc
ggatatggtccagcgttgtaaagcggtgggggtcgatatttatctggatgcggtgatcaaccatatggcagcacaagatcgctattttccagaagt
accttacagcagtaatgatttttcacagttgcacgggcgatatcgattattccaaccgctggtcgattcaaaattgcgatctggttgggctgaacgat
ctcaaaaccgagtcagaatacgttcggcagaaaattgcagactatatgaacgatgcgctcagtctgggcgtggcggggtttcggattgatgccg
ccaagcatatcccggccggcgacatcgcggcgatcaagagcaagctcaacggcagcccgtatatctatcaggaggttatcggggcggcagg
ggagccggtacaaaccagcgagtacacgtatattggagacgtgacggaatttaacttcgcccggaccatcgggcctaaatttaagcaaggtaat
attaaagacctgcaggggattggttcgtggagcggctggctgagcagcgacgatgcggtgacctttgtgaccaaccatgacgaagaacgcca
taaccctggccaggttctcagccatcaggactttggcaatctgtatttcctcggtaacgtgtttactctggcgtatccttacggctacccaaaagtga
tgtcggggtactacttcagtaattttgatgccgggccaccatcgacaggggtacattctggtaatgcgtgtggctttgatggcggtgattgggtctg
cgaacacaaatggcgtggtgtaggcaacatggtggcgttcgcaaccgacagcgccgagtggcaggtcaatgactggtgggacgatggtt
acaatcaggtggcgttcggtcgtggcaacaaggctgggcttgtggtgatcaatcgagatgacaataatggcatcaatcagagtttccagacgggaat
gcccgctggcgagtattgtgacatcattgccggtgatttcgacacccagagcggtcattgcagcgctacgacgatcaccgtcgacagtcaggg
gtatgcacattttactgtcggtagtcatcaggccgctgcgattcacattggcgcgaaactcggctccgtgtgccaggactgtggcggcacggcc
gcagagacaaaagtctgctttgacaatgcacaaaactttagccaaccgtatttgcattactggaatgtcaatgcggatcaggccgtagcgaatgc
aacctggccgggcgtcgcgatgacggctgaaaatggcggttactgctacgattttggtgtcggtctcaattcacttcaggtaattttcagcgataa
cggcgccagccaaaccgctgatctgaccgccagcagtccgacgttgtgttaccagaacggaacgtggcgtgacagtgacttctgtcagagta
gcaatgtgggcaacgagagttggtatttccgtggaacctcaaacggttggggcgtgagcgcactcacttatgaggctgcgacaggcctgtaca
ctacggtgcagagctttaacggggaggagtcgcccgcacgctttaaaattgatgatggcaactggagtgagtcgtatccaagtgctgattatcaa
gtcggtgattatgccacctacacgatcacgtttgacagccagacgaaggccatcaccgtgacttcgcagtaa

SEQ ID NO: 194

Met Lys Phe Lys Lys Ser Leu Ser Ala Gly Leu Leu Leu Phe Gly Gly Leu Ser Gly Val Thr Pro Ser Val Ala Ala Glu Val Pro Arg Thr Ala Phe Val His Leu Phe Glu Trp Ser Trp Pro Asp Ile Ala Thr

FIGURE 16EEEE

Glu Cys Glu Thr Phe Leu Gly Pro Lys Gly Phe Ser Ala Val Gln Val Ser Pro Pro Gln Lys Ser Val Ser Asn Ala Ala Trp Trp Ala Arg Tyr Gln Pro Val Ser Tyr Ser Phe Glu Gly Arg Ser Gly Thr Arg Ala Gln Phe Ala Asp Met Val Gln Arg Cys Lys Ala Val Gly Val Asp Ile Tyr Leu Asp Ala Val Ile Asn His Met Ala Ala Gln Asp Arg Tyr Phe Pro Glu Val Pro Tyr Ser Ser Asn Asp Phe His Ser Cys Thr Gly Asp Ile Asp Tyr Ser Asn Arg Trp Ser Ile Gln Asn Cys Asp Leu Val Gly Leu Asn Asp Leu Lys Thr Glu Ser Glu Tyr Val Arg Gln Lys Ile Ala Asp Tyr Met Asn Asp Ala Leu Ser Leu Gly Val Ala Gly Phe Arg Ile Asp Ala Ala Lys His Ile Pro Ala Gly Asp Ile Ala Ala Ile Lys Ser Lys Leu Asn Gly Ser Pro Tyr Ile Tyr Gln Glu Val Ile Gly Ala Ala Gly Glu Pro Val Gln Thr Ser Glu Tyr Thr Tyr Ile Gly Asp Val Thr Glu Phe Asn Phe Ala Arg Thr Ile Gly Pro Lys Phe Lys Gln Gly Asn Ile Lys Asp Leu Gln Gly Ile Gly Ser Trp Ser Gly Trp Leu Ser Ser Asp Asp Ala Val Thr Phe Val Thr Asn His Asp Glu Glu Arg His Asn Pro Gly Gln Val Leu Ser His Gln Asp Phe Gly Asn Leu Tyr Phe Leu Gly Asn Val Phe Thr Leu Ala Tyr Pro Tyr Gly Tyr Pro Lys Val Met Ser Gly Tyr Tyr Phe Ser Asn Phe Asp Ala Gly Pro Pro Ser Thr Gly Val His Ser Gly Asn Ala Cys Gly Phe Asp Gly Gly Asp Trp Val Cys Glu His Lys Trp Arg Gly Val Ala Asn Met Val Ala Phe Arg Asn His Thr Ala Ala Gln Trp Gln Val Thr Asp Trp Trp Asp Asp Gly Tyr Asn Gln Val Ala Phe Gly Arg Gly Gly Leu Gly Phe Val Val Ile Asn Arg Asp Asp Asn Lys Gly Ile Asn Gln Ser Phe Gln Thr Gly Met Pro Ala Gly Glu Tyr Cys Asp Ile Ile Ala Gly Asp Phe Asp Thr Gln Ser Gly His Cys Ser Ala Thr Thr Ile Thr Val Asp Ser Gln Gly Tyr Ala His Phe Thr Val Gly Ser His Gln Ala Ala Ala Ile His Ile Gly Ala Lys Leu Gly Ser Val Cys Gln Asp Cys Gly Gly Thr Ala Ala Glu Thr Lys Val Cys Phe Asp Asn Ala Gln Asn Phe Ser Gln Pro Tyr Leu His Tyr Trp Asn Val Asn Ala Asp Gln Ala Val Ala Asn Ala Thr Trp Pro Gly Val Ala Met Thr Ala Glu Asn Gly Gly Tyr Cys Tyr Asp Phe Gly Val Gly Leu Asn Ser Leu Gln Val Ile Phe Ser Asp Asn Gly Ala Ser Gln Thr Ala Asp Leu Thr Ala Ser Ser Pro Thr Leu Cys Tyr Gln Asn Gly Thr Trp Arg Asp Ser Asp Phe Cys Gln Ser Ser Asn Val Gly Asn Glu Ser Trp Tyr Phe Arg Gly Thr Ser Asn Gly Trp Gly Val Ser Ala Leu Thr Tyr Glu Ala Ala Thr Gly Leu Tyr Thr Thr Val Gln Ser Phe Asn Gly Glu Glu Ser Pro Ala Arg Phe Lys Ile Asp Asp Gly Asn Trp Ser Glu Ser Tyr Pro Ser Ala Asp Tyr Gln Val Gly Asp Tyr Ala Thr Tyr Thr Ile Thr Phe Asp Ser Gln Thr Lys Ala Ile Thr Val Thr Ser Gln

SEQ ID NO: 195

atgctgacagaccgtttctttgatggcgatacatcaaacaacgacccttacaaccagaactacgatgctaaaaacgaccggggaacttatcagg
gcggcgattttaaaggaatcacgcaaaaattggattatctcgataagctaggcgtgaacacaatctggatcagcccgatcgtggaaaatatcaag
catgatgtccgttatgacaactctgaagggcattcatactatgcttaccacggctactgggcaagcaacttcggtgcgttaaacccacacttcggt
acaatggaagatttccatacactgattgacgctgcccatgaaaaaggcatcaagatcatggttgacgtagtattaaaccacactggttatggctta
aaagatatcaacggagaagtttccaatcctccagccggttacccaactgacgcagaacgcagcacatatagcagcctgcttcgccagggttca
aatgtcggctctgatgaggttgttggcgaattagctggcctacctgacttaaaaacagaagacccgcagtccgccagacaatcatcgactggc
aaacagactggatcacgaaagctactacagctaaaggaaacacaattgactacttccgtgtcgacactgtgaagcacgttgaagacgcaacat
ggatggcattcaaaaatgacctcactgataaaatgcgcgacacacaaaatgatcgggtaagtttgggggagcaaggtgccaatttacaaaacttggat
accttgaaacaggtatgatggactcactgcttgacttcgacttcaaaggcattgcgcacgatttcgtgaacggcaagcttaaggcagcaaacgat
gccctgactgcccgcaacggtaaaattgacaacacagctactttaggttcattccttggaagccatgacgaagatggtttcctatttaaagaagga
aatgacaaaggcaagcttaaggttgctgcttccctgcaagcaacatcaaaaggccagccggtcatctattatggtgaagagcttggtcaaagtg
gagcaaacaactatccgcaatacgataaccgttatgacctggcatgggacaaagttgaaaacaacgacgtccttgagcactacactaaggtcct
gaacttcagaagcgctcattcagaagtgttcgctaaaggtgaacgcgcaacaattggcggttctgacgctgataaattcttactttttgctctgtaaa
aatggaaacgaagctgcttacgtcggcttgaacgttgctgacacagcaaaagacgtaacactgactgtttctgcaggtgcagtcgtaactgacc
actatgcagataaaacttatactgcttcagaagctggagaaatcacattgacgatcccggcaaaagctgatggcggtactgttttactaacggttg
aaggcggagaaatcacagctgctaaagcggcaagcgaaggcgacggcacagttgagccagtccctgcgaaccacatccgcattcactacaa
ccgtacagacaacaactatgaaaactacggtgcatggctgtggaacgatgtagcctcccccttctgccaactggccgactggcgctacaatgtttg
aaaaaacagacagctacggtgcatacatcgacgtaccacttaaagagggcgctaagaacatcggcttcctcgttatggatgtaacaaaaggtga
tcagggtaaagacggcggcgacaaaggttttacgatctcatcacctgaaatgaacgaaatttggatcaagcaaggttctgacaaggtgtacactt
acgagccagttgatcttccggcgaacactgtccgcgtccactatgtacgtgacaacgcagactacgaaaacttcggtatctggaactggggcga
tgtaacagcaccttccgaaaactggcctacaggcgcagcgaaattcgatggtacagaccgttacggtgcgtatgtcgacattacgctaaaagaa

FIGURE 16FFFF

ggcgcaaagaacattggaatgattgctcttaacactgcaaatggagagaaagacggcggagataaatccttcaaccttctggataaatataatcg
catttggattaaacaaggtgatgacaatgtctacgtttctccatactgggagcaggcaacaggaatcaccaatgcagaggtaatctctgaagata
cgattctattaggcttcacaatgactgacggcttaacacctgaatctttaaaaggaggtcttgtaattaaagattcaactggtgctgaagttgccatc
gaaagtgctgaaatcacaagcgcaacctctgtaaaagtaaaagcaacattcgatttagaaaagcttccattatccatcacatacgcaggcagaac
agtttcagcttcaactggctggagaatgcttgatgaaatgtacgcttatgatggaaacgaccttggtgcgacttacaaggacggagcagcgacg
cttaaattatgggctccgaaagcgagcaaggtaaccgctaacttctttgataaaaataatgccgctgaaaaaatcggcagcgtcgagttaacgaa
gggtgaaaaaggagtctggtcagctatggttgctcctggcgacctgaacgtaaccgatcttgaaggttatttttaccagtatgatgtaacaaatga
cggtataactcgccaggtgttagatccttatgcaaaatcaatggcagcctttactgtgaatacagaaggcaatgctggtcctgacggggacactg
ttggcaaggcggcaattcaaaaagcttctcgagagtacttctag

SEQ ID NO: 196

Met Leu Thr Asp Arg Phe Phe Asp Gly Asp Thr Ser Asn Asn Asp Pro Tyr Asn Gln Asn Tyr Asp
Ala Lys Asn Asp Arg Gly Thr Tyr Gln Gly Gly Asp Phe Lys Gly Ile Thr Gln Lys Leu Asp Tyr Leu
Asp Lys Leu Gly Val Asn Thr Ile Trp Ile Ser Pro Ile Val Glu Asn Ile Lys His Asp Val Arg Tyr Asp
Asn Ser Glu Gly His Ser Tyr Tyr Ala Tyr His Gly Tyr Trp Ala Ser Asn Phe Gly Ala Leu Asn Pro
His Phe Gly Thr Met Glu Asp Phe His Thr Leu Ile Asp Ala Ala His Glu Lys Gly Ile Lys Ile Met
Val Asp Val Val Leu Asn His Thr Gly Tyr Gly Leu Lys Asp Ile Asn Gly Glu Val Ser Asn Pro Pro
Ala Gly Tyr Pro Thr Asp Ala Glu Arg Ser Thr Tyr Ser Ser Leu Leu Arg Gln Gly Ser Asn Val Gly
Ser Asp Glu Val Val Gly Glu Leu Ala Gly Leu Pro Asp Leu Lys Thr Glu Asp Pro Ala Val Arg Gln
Thr Ile Ile Asp Trp Gln Thr Asp Trp Ile Thr Lys Ala Thr Thr Ala Lys Gly Asn Thr Ile Asp Tyr Phe
Arg Val Asp Thr Val Lys His Val Glu Asp Ala Thr Trp Met Ala Phe Lys Asn Asp Leu Thr Glu
Lys Met Pro Thr His Lys Met Ile Gly Glu Ala Trp Gly Ala Ser Ala Asn Asn Gln Leu Gly Tyr Leu
Glu Thr Gly Met Met Asp Ser Leu Leu Asp Phe Asp Phe Lys Gly Ile Ala His Asp Phe Val Asn Gly
Lys Leu Lys Ala Ala Asn Asp Ala Leu Thr Ala Arg Asn Gly Lys Ile Asp Asn Thr Ala Thr Leu Gly
Ser Phe Leu Gly Ser His Asp Glu Asp Gly Phe Leu Phe Lys Glu Gly Asn Asp Lys Gly Lys Leu
Lys Val Ala Ala Ser Leu Gln Ala Thr Ser Lys Gly Gln Pro Val Ile Tyr Tyr Gly Glu Glu Leu Gly
Gln Ser Gly Ala Asn Asn Tyr Pro Gln Tyr Asp Asn Arg Tyr Asp Leu Ala Trp Asp Lys Val Glu
Asn Asn Asp Val Leu Glu His Tyr Thr Lys Val Leu Asn Phe Arg Ser Ala His Ser Glu Val Phe Ala
Lys Gly Glu Arg Ala Thr Ile Gly Gly Ser Asp Ala Asp Lys Phe Leu Leu Phe Ala Arg Lys Asn Gly
Asn Glu Ala Ala Tyr Val Gly Leu Asn Val Ala Asp Thr Ala Lys Asp Val Thr Leu Thr Val Ser Ala
Gly Ala Val Val Thr Asp His Tyr Ala Asp Lys Thr Tyr Thr Ala Ser Glu Ala Gly Glu Ile Thr Leu
Thr Ile Pro Ala Lys Ala Asp Gly Gly Thr Val Leu Leu Thr Val Glu Gly Gly Glu Ile Thr Ala Ala
Lys Ala Ala Ser Glu Gly Asp Gly Thr Val Glu Pro Val Pro Ala Asn His Ile Arg Ile His Tyr Asn
Arg Thr Asp Asn Asn Tyr Glu Asn Tyr Gly Ala Trp Leu Trp Asn Asp Val Ala Ser Pro Ser Ala Asn
Trp Pro Thr Gly Ala Thr Met Phe Glu Lys Thr Asp Ser Tyr Gly Ala Tyr Ile Asp Val Pro Leu Lys
Glu Gly Ala Lys Asn Ile Gly Phe Leu Val Met Asp Val Thr Lys Gly Asp Gln Gly Lys Asp Gly Gly
Asp Lys Gly Phe Thr Ile Ser Ser Pro Glu Met Asn Glu Ile Trp Ile Lys Glu Gly Ser Asp Lys Val
Tyr Thr Tyr Glu Pro Val Asp Leu Pro Ala Asn Thr Val Arg Val His Tyr Val Arg Asp Asn Ala Asp
Tyr Glu Asn Phe Gly Ile Trp Asn Trp Gly Asp Val Thr Ala Pro Ser Glu Asn Trp Pro Thr Gly Ala
Ala Lys Phe Asp Gly Thr Asp Arg Tyr Gly Ala Tyr Val Asp Ile Thr Leu Lys Glu Gly Ala Lys Asn
Ile Gly Met Ile Ala Leu Asn Thr Ala Asn Gly Glu Lys Asp Gly Gly Asp Lys Ser Phe Asn Leu Leu
Asp Lys Tyr Asn Arg Ile Trp Ile Lys Gln Gly Asp Asp Asn Val Tyr Val Ser Pro Tyr Trp Glu Gln
Ala Thr Gly Ile Thr Asn Ala Glu Val Ile Ser Glu Asp Thr Ile Leu Leu Gly Phe Thr Met Thr Asp
Gly Leu Thr Pro Glu Ser Leu Lys Gly Gly Leu Val Ile Lys Asp Ser Thr Gly Ala Glu Val Ala Ile
Glu Ser Ala Glu Ile Thr Ser Ala Thr Ser Val Lys Val Lys Ala Thr Phe Asp Leu Glu Lys Leu Pro
Leu Ser Ile Thr Tyr Ala Gly Arg Thr Val Ser Ala Ser Thr Gly Trp Arg Met Leu Asp Glu Met Tyr
Ala Tyr Asp Gly Asn Asp Leu Gly Ala Thr Tyr Lys Asp Gly Ala Ala Thr Leu Lys Leu Trp Ala Pro
Lys Ala Ser Lys Val Thr Ala Asn Phe Phe Asp Lys Asn Asn Ala Ala Glu Lys Ile Gly Ser Val Glu
Leu Thr Lys Gly Glu Lys Gly Val Trp Ser Ala Met Val Ala Pro Gly Asp Leu Asn Val Thr Asp Leu
Glu Gly Tyr Phe Tyr Gln Tyr Asp Val Thr Asn Asp Gly Ile Thr Arg Gln Val Leu Asp Pro Tyr Ala

FIGURE 16GGGG

Lys Ser Met Ala Ala Phe Thr Val Asn Thr Glu Gly Asn Ala Gly Pro Asp Gly Asp Thr Val Gly Lys Ala Ala Ile Gln Lys Ala Ser Arg Glu Tyr Phe

SEQ ID NO: 197

atgaaaccgtcaaaattcgtttttctctctgctgccatcgcttgcagcctctccagtaccgccaatgctgacgccattttgcatgcatttaactggaag
tactccgacgtcacgcaaaacgcctcgcaaatcgcggcggcgggttataaaaaagtgctgatttcgccagcactgaaatcgagtggcaatgaa
tggtgggcacgttatcaaccgcaagatctgcgcgtgatcgattccccacttggcaacaaaagtgacttaaaatccatgattgatgctctgaaggc
ggtcggcgttgatgtgtatgccgatgtggtgcttaaccatatggccaatgaaacatggaagcgtgaagacttaaattaccctggcagtgaagtgc
tgcaacaatacgcagctaacaccagttattatgcggaccaaacgctttttggcaatttaacggaaaacctattctctggctttgacttccacccaga
aggctgtattagcgattggaatgatgccggcaatgttcagtactggcgtctttgtggcggtgctggtgaccgagggctgccagacttagatccga
acaactgggtggtgtcacagcaacgtttgtatttgaatgcgctaaaaggtttaggtgtgaaaggcttccgcattgatgcggttaaacacatgagcc
aatatcaaatcgaccagattttcactgcagagattaccgccggaatgcacgtgtttggtgaagtgatcaccagtggtggcaaaggcgactccag
ctatgagaacttcttagcgccttatctcaacgccaccaaccattcggcttacgatttcccactgtttgcctctattcgcaacgccttctcctacagcgg
tggcatgaacatgcttcatgatccacaagcctatggccaagggcttgaaaacgcacgttcaattacctttaccatcacgcacgacatcccaacga
acgacggtttccgttatcaaatcatggatccgaaagatgaagagctggcttacgcttatatcctcggtaaagatggcggcacacctctgatttaca
gcgacaacttacctgataacgaagatcgtgataatcgccgttgggaaggtgtttggaaccgtgacctgatgaagaacatgttgcgcttccataac
caaatgcaagggcaagagatgacgatgctgtacagcgaccaatgtctactgatgtttaagcgcggtaaacaaggggtggtcggcattaataaat
gcggtgaagagcgttctcataccgttgacacctatcagcatgagttcaactggtatcagccttacacagatacactcactggcgtgactgaaacc
gtgagttcgcgttaccacaccttccgaattccagctcgcagcgcgcgcatgtacatgctctaa

SEQ ID NO: 198

Met Lys Pro Ser Lys Phe Val Phe Leu Ser Ala Ala Ile Ala Cys Ser Leu Ser Ser Thr Ala Asn Ala Asp Ala Ile Leu His Ala Phe Asn Trp Lys Tyr Ser Asp Val Thr Gln Asn Ala Ser Gln Ile Ala Ala Ala Gly Tyr Lys Lys Val Leu Ile Ser Pro Ala Leu Lys Ser Ser Gly Asn Glu Trp Trp Ala Arg Tyr Gln Pro Gln Asp Leu Arg Val Ile Asp Ser Pro Leu Gly Asn Lys Ser Asp Leu Lys Ser Met Ile Asp Ala Leu Lys Ala Val Gly Val Asp Val Tyr Ala Asp Val Val Leu Asn His Met Ala Asn Glu Thr Trp Lys Arg Glu Asp Leu Asn Tyr Pro Gly Ser Glu Val Leu Gln Gln Tyr Ala Ala Asn Thr Ser Tyr Tyr Ala Asp Gln Thr Leu Phe Gly Asn Leu Thr Glu Asn Leu Phe Ser Gly Phe Asp Phe His Pro Glu Gly Cys Ile Ser Asp Trp Asn Asp Ala Gly Asn Val Gln Tyr Trp Arg Leu Cys Gly Gly Ala Gly Asp Arg Gly Leu Pro Asp Leu Asp Pro Asn Asn Trp Val Val Ser Gln Gln Arg Leu Tyr Leu Asn Ala Leu Lys Gly Leu Gly Val Lys Gly Phe Arg Ile Asp Ala Val Lys His Met Ser Gln Tyr Gln Ile Asp Gln Ile Phe Thr Ala Glu Ile Thr Ala Gly Met His Val Phe Gly Glu Val Ile Thr Ser Gly Gly Lys Gly Asp Ser Ser Tyr Glu Asn Phe Leu Ala Pro Tyr Leu Asn Ala Thr Asn His Ser Ala Tyr Asp Phe Pro Leu Phe Ala Ser Ile Arg Asn Ala Phe Ser Tyr Ser Gly Gly Met Asn Met Leu His Asp Pro Gln Ala Tyr Gly Gln Gly Leu Glu Asn Ala Arg Ser Ile Thr Phe Thr Ile Thr His Asp Ile Pro Thr Asn Asp Gly Phe Arg Tyr Gln Ile Met Asp Pro Lys Asp Glu Glu Leu Ala Tyr Ala Tyr Ile Leu Gly Lys Asp Gly Gly Thr Pro Leu Ile Tyr Ser Asp Asn Leu Pro Asp Asn Glu Asp Arg Asp Asn Arg Arg Trp Glu Gly Val Trp Asn Arg Asp Leu Met Lys Asn Met Leu Arg Phe His Asn Gln Met Gln Gly Gln Glu Met Thr Met Leu Tyr Ser Asp Gln Cys Leu Leu Met Phe Lys Arg Gly Lys Gln Gly Val Val Gly Ile Asn Lys Cys Gly Glu Glu Arg Ser His Thr Val Asp Thr Tyr Gln His Glu Phe Asn Trp Tyr Gln Pro Tyr Thr Asp Thr Leu Thr Gly Val Thr Glu Thr Val Ser Ser Arg Tyr His Thr Phe Arg Ile Pro Ala Arg Ser Ala Arg Met Tyr Met Leu

SEQ ID NO: 199

gtgagtttgaccaaaaaggctcagtacgaaccaaatacggcaccaaggctcagtacatctctgcaatcaatgccgcgcacaacaacaatatcca
aatttacggcgatgttgtgtttaaccaccgaggtggtgctgatgggaagtcgtgggtcgataccaagcgcgttgattgggacaaccgcaatattg
aactgggcgacaaatggattgaagcttgggttgagtttaattttcctggccgcaacgacaaatactcgaacttccattggacttggtatcactttgac
ggtgttgactgggatgacgccggcaaagaaaaagcgatctttaaattcaaaggcgaaggaaaagcatgggattgggaagtcagctctgaaaa
aggcaattacgactacctaa

FIGURE 16HHHH

SEQ ID NO: 200

Val Ser Leu Thr Lys Lys Ala Gln Tyr Glu Pro Asn Thr Ala Pro Arg Leu Ser Thr Ser Leu Gln Ser Met Pro Arg Thr Thr Thr Ile Ser Lys Phe Thr Ala Met Leu Cys Leu Thr Thr Glu Val Val Leu Met Gly Ser Arg Gly Ser Ile Pro Ser Ala Leu Ile Gly Thr Thr Ala Ile Leu Asn Trp Ala Thr Asn Gly Leu Lys Leu Gly Leu Ser Leu Ile Phe Leu Ala Ala Thr Thr Asn Thr Arg Thr Ser Ile Gly Leu Gly Ile Thr Leu Thr Val Leu Thr Gly Met Thr Pro Ala Lys Lys Lys Arg Ser Leu Asn Ser Lys Ala Lys Glu Lys His Gly Ile Gly Lys Ser Ala Leu Lys Lys Ala Ile Thr Thr Thr

SEQ ID NO: 201

atgacagccaaggctgatgacttacgcatttaccagatcatggtggaaagctttgtggatggcgataaacaggtcggccatggcaccggctacg
gtaccagccatcacaaaggcgatctgcaagggatcattgactcgctggattacattcaatcgctgggcgtcaatgccatttggctaacgccgattt
ttgaatctattccggtggagggacaagaccattgggcggacaggcttgatgctacaggctactttgccagtgactatttcaagatagaccctcgct
ttggcacgttagaacaagcccgtgagctggtggaaaaggcacacgcgaaaggcttgtatgtcttctttgatggagtatttggtcaccataaaggc
aatgtggtgccatcaccacaaggtagactgcctgtcggtgaaaataacccggtcagctacccagagagcctggcgttttacgaagaagtcgcc
agttactgggtgaaagagttaaagattgatggctggcgtctggatcaagcctatcaagtgccgaccgatgcatggaaagcgatccgtcagagc
gttgatgaagcgtcacagtccgtaacttatgtgaataacaaaggggaaaccgtccatcctttgggttacatggtggctgaaatttggaataacgaa
cgttacatcacagaaaccggttacggcaaagaaggcgatccggcgttgtgctcggcttttgattttccgatgcgtttccgagtggtcgaaaccttt
gcggttaacgaaagtggtgtcagccgaaaaggcggcgaatggttgaatgacggcatgtcactgcacagtcagtatccggatcatgccaagcct
aatttaatgttgggcaaccatgatgtggtgcgctttggggatctgctgcaacgtggcggtattgcgtcaccagaacaaccgcaatactggcagcg
tcataaagcggcgatgtctttcttagcagcgtataccggcccaattaccttgtattacggtgaagaaattggcgatcaggttgacggctttgctaaa
aaaatcaaagaagattgtgccgttattggtttgtgtgatgaccacgtggcgcgcaccagtgcgaagattgatggcgtgacggcgtcactgaatg
cacagcagtctgaactcaaagtatatgtctcttcattgatgacattacgtcagcaacatcctgcgttatcacaaggggaacgtactaatgtgatggc
gacagagacagtatacgtagaccataaacaggcagacaatgaagccctgttgtacatggtgagtacgactgataacgcggagtcagtcaccctt
gaagggcaaagcgattggttcacaaggtgtgctgattgatttgttaacgaacgagcgttttatgcccaataatggggagtatgccattccattaac
gggctttggcgcacgattcctcaagattgacactccgacagcggcgggtgtgatggcgcaatctgctgcctcggtatcgctagtaggtgaagg
gatcatggcccaatgtgataccccaaccgttgaaggcaccggtccggtagcagaaaccttgtacgtggttggcgattttgccgatgctggttgga
agcaaaagccgcagcgcgcgtatcaatacaaaggcaagcacaatggcagcaacttgtatcaagtggttgtcgatgaaaaagcgggcgcctac
aagatgcaatacgccacgaaagattggagcccacagtttactgcagacggtatggcattgaagccgggtaccgcaaagtcgctcatagcgggt
ggctacggtaaagacaccgccgtgacgttgccggaatccggtaagtatgtgtggagcttaacattcagtgatcttggcgagccggagcaaatc
atggtgtctaagtgtcagtaa

SEQ ID NO: 202

Met Thr Ala Lys Ala Asp Asp Leu Arg Ile Tyr Gln Ile Met Val Glu Ser Phe Val Asp Gly Asp Lys Gln Val Gly His Gly Thr Gly Tyr Gly Thr Ser His His Lys Gly Asp Leu Gln Gly Ile Ile Asp Ser Leu Asp Tyr Ile Gln Ser Leu Gly Val Asn Ala Ile Trp Leu Thr Pro Ile Phe Glu Ser Ile Pro Val Glu Gly Gln Asp His Trp Ala Asp Arg Leu Asp Ala Thr Gly Tyr Phe Ala Ser Asp Tyr Phe Lys Ile Asp Pro Arg Phe Gly Thr Leu Glu Gln Ala Arg Glu Leu Val Glu Lys Ala His Ala Lys Gly Leu Tyr Val Phe Phe Asp Gly Val Ile Gly His His Lys Gly Asn Val Val Pro Ser Pro Gln Gly Arg Leu Pro Val Gly Glu Asn Asn Pro Val Ser Tyr Pro Glu Ser Leu Ala Phe Tyr Glu Glu Val Ala Ser Tyr Trp Val Lys Glu Leu Lys Ile Asp Gly Trp Arg Leu Asp Gln Ala Tyr Gln Val Pro Thr Asp Ala Trp Lys Ala Ile Arg Gln Ser Val Asp Glu Ala Ser Gln Ser Val Thr Tyr Val Asn Asn Lys Gly Glu Thr Val His Pro Leu Gly Tyr Met Val Ala Glu Ile Trp Asn Asn Glu Arg Tyr Ile Thr Glu Thr Gly Tyr Gly Lys Glu Gly Asp Pro Ala Leu Cys Ser Ala Phe Asp Phe Pro Met Arg Phe Arg Val Val Glu Thr Phe Ala Val Asn Glu Ser Gly Val Ser Arg Lys Gly Gly Glu Trp Leu Asn Asp Gly Met Ser Leu His Ser Gln Tyr Pro Asp His Ala Lys Pro Asn Leu Met Leu Gly Asn His Asp Val Val Arg Phe Gly Asp Leu Leu Gln Arg Gly Gly Ile Ala Ser Pro Glu Gln Pro Gln Tyr Trp Gln Arg His Lys Ala Ala Met Ser Phe Leu Ala Ala Tyr Thr Gly Pro Ile Thr Leu Tyr Tyr Gly Glu Glu Ile Gly Asp Gln Val Asp Gly Phe Ala Lys Lys Ile Lys Glu Asp Cys Ala Val Ile Gly Leu Cys Asp Asp His Val Ala Arg Thr Ser Ala Lys Ile Asp Gly Val Thr Ala Ser Leu Asn Ala Gln Gln Ser Glu Leu Lys Val Tyr Val Ser Ser Leu Met Thr Leu Arg Gln Gln His Pro Ala Leu Ser Gln Gly Glu Arg Thr Asn Val Met Ala Thr Glu

FIGURE 16IIII

Thr Val Tyr Val Asp His Lys Gln Ala Asp Asn Glu Ala Leu Leu Tyr Met Val Ser Thr Thr Asp Asn Ala Glu Ser Val Thr Leu Lys Gly Lys Ala Ile Gly Ser Gln Gly Val Leu Ile Asp Leu Leu Thr Asn Glu Arg Phe Met Pro Asn Asn Gly Glu Tyr Ala Ile Pro Leu Thr Gly Phe Gly Ala Arg Phe Leu Lys Ile Asp Thr Pro Thr Ala Ala Gly Val Met Ala Gln Ser Ala Ala Ser Val Ser Leu Val Gly Glu Gly Ile Met Ala Gln Cys Asp Thr Pro Thr Val Glu Gly Thr Gly Pro Val Ala Glu Thr Leu Tyr Val Val Gly Asp Phe Ala Asp Ala Gly Trp Lys Gln Lys Pro Gln Arg Ala Tyr Gln Tyr Lys Gly Lys His Asn Gly Ser Asn Leu Tyr Gln Val Val Val Asp Glu Lys Ala Gly Ala Tyr Lys Met Gln Tyr Ala Thr Lys Asp Trp Ser Pro Gln Phe Thr Ala Asp Gly Met Ala Leu Lys Pro Gly Thr Ala Lys Ser Leu Ile Ala Gly Gly Tyr Gly Lys Asp Thr Ala Val Thr Leu Pro Glu Ser Gly Lys Tyr Val Trp Ser Leu Thr Phe Ser Asp Leu Gly Glu Pro Glu Gln Ile

SEQ ID NO: 203

atgaagatgaagtcccgggcgtggttgttaggtagtgcagtggccatggcgttggcctcttcggcagccaatgccggtgtcatggttcacctgtt
ccagtggaagtacaatgacatcgccaacgagtgcgaaaaggtgctcggtcccaaagggtatgaagcagtgcagatcacgccgcctgctgaa
cacctgcaaggctcctcctggtgggtggtctatcagcccgtcagctacaagaacttcacttctctgggcggtaacgaggccgaactcaaaagca
tgatcgcccgttgcaaggccgccggggtcaagatttacgccgatgcggtattcaaccagctggctggtggatcaggcgtcggtacaggtggta
gcagctacaatgccggcagcttcagctatccccaatttggctacaacgatttccatcacgctgggagcctcaccaactatgccgaccgcaacaa
tgtgcaaaacggtgccctgctggggctgccggatctggataccggctctgcctatgtgcaggatcagctggctacctatatgaagaccctgagt
ggctggggtgtggcaggttttcgtcttgatgcagcaaagcatatgagcgttgccgatctctcggccatcgtcagcaaggcgggcaatccttttgt
ctactccgaggtgattggtgccacgggtgaaccaatccagccgggcgaatataccggcattggtgccgtgaccgaatttaaatacggcaccga
tctggcctccaacttcaaggggcagatcaagaatctcaagagcatgggcgagagctggggtctgcttgcgtcgaacaaggctgaagtctttgtg
gtcaaccatgaccgtgagcggggacatggcggtggcggtatgctgacctacaaggatggtgccctctacaatctggccaacatcttcatgctg
gcctggccctatggcgcctatccccaggtgatgtccggctatgatttcggcaccaatacccgatattggtgggccgagcgctacccccttgttcttcc
ggctctagctggaactgcgaacaccgctggagcaacatcgccaacatggtctcgttccacaatgccgcccaaggcacgtccatgaccaactg
gtgggataatggtaataaccagatcgcctttggtcgcggcgccaaggcctttgtggtgatcaacaatgaatcttccactctgagcaagagcctgc
agacgggtctgccagccgggggagtactgcaacattctggccggtgatgccctgtgcagcggcagcaccatcaaggtggatgccagcggtat
ggccaccttcaacgtggcagggatgaaggcggcagcgatccatatcaatgccaagcccgatagcaccagcagtggcagctcaggctcttcct
ctggctcttcttcctctgccaccagtaacaagtttgccagcatgaatctgcggggcaccaacaatggctgggccagcaccgccatgacagtgga
tgccaaccgtgtctggtcggcggatgtcacctttaccggggccgcggatgccaatggtgcccagcgcttcaagtttgatgtctatggcaactgg
acagagagctatggcgatacacaagccgatggcattgccgacaaggggagcgccaaggacatctatttcaatggtgtgggcaagtatcgtgtc
tcgctcaaggagagcgacatgagctacaccctgacccagctctccagcaatcaggcaccggtggcggccatcacccccaagacactctccgt
caagctgggtgactcagtggtgttcgatgcctccggctccaccgatgatgtgggtgtcactggctacagctggtctaccggtggcagtgccaag
accgaaactgtgctgtttgatgctctgggtaccaagaccattaccgt gtggccgatgccgatggcttgacctccaaggccagtgccaccg
tcaccgtcaccgatggcagcgtggcttataacagcaactttgccagcctgaacttccgtggcactcccaacagttggggcgcggcagccatga
cgctggtggcagacaacacctgggaggcaacggtcaacttcgatggtcaggccaatcagcgcttcaagttcgatatcaagggtgactggagc
cagaactatggtgatagcaacaaggatggggtggccgaacgtaccggtgccgatatttacaccactgtgaccggtcaatataaggtgcaattta
acgactccacttgaagtacaccctgaccaagctggccgatggcagcgccaccagctatagcgcgaacttggccagcctctaccfgcgtggca
ccccgaacagctgggcaaccacgccatgaagctggtggccaataacagctggcaggccgaggtgaccttcaccggcaagggcgatgcca
ctggtgcccaacgcttcaagttcgacgtcaagggtgactggagccagaactacggtgacagcaacatggacgggactgccgaacggactgg
tggcgatatcaccagtgccgtggtgggcacctatctggtgacctttaatgacagcacactgaaatacaccctgaccgccaaataa

SEQ ID NO: 204

Met Lys Met Lys Ser Arg Ala Trp Leu Leu Gly Ser Ala Val Ala Met Ala Leu Ala Ser Ser Ala Ala Asn Ala Gly Val Met Val His Leu Phe Gln Trp Lys Tyr Asn Asp Ile Ala Asn Glu Cys Glu Lys Val Leu Gly Pro Lys Gly Tyr Glu Ala Val Gln Ile Thr Pro Pro Ala Glu His Leu Gln Gly Ser Ser Trp Trp Val Val Tyr Gln Pro Val Ser Tyr Lys Asn Phe Thr Ser Leu Gly Gly Asn Glu Ala Glu Leu Lys Ser Met Ile Ala Arg Cys Lys Ala Ala Gly Val Lys Ile Tyr Ala Asp Ala Val Phe Asn Gln Leu Ala Gly Gly Ser Gly Val Gly Thr Gly Gly Ser Ser Tyr Asn Ala Gly Ser Phe Ser Tyr Pro Gln Phe Gly Tyr Asn Asp Phe His His Ala Gly Ser Leu Thr Asn Tyr Ala Asp Arg Asn Asn Val Gln Asn Gly Ala Leu Leu Gly Leu Pro Asp Leu Asp Thr Gly Ser Ala Tyr Val Gln Asp Gln Leu Ala Thr Tyr Met

FIGURE 16JJJ

Lys Thr Leu Ser Gly Trp Gly Val Ala Gly Phe Arg Leu Asp Ala Ala Lys His Met Ser Val Ala Asp Leu Ser Ala Ile Val Ser Lys Ala Gly Asn Pro Phe Val Tyr Ser Glu Val Ile Gly Ala Thr Gly Glu Pro Ile Gln Pro Gly Glu Tyr Thr Gly Ile Gly Ala Val Thr Glu Phe Lys Tyr Gly Thr Asp Leu Ala Ser Asn Phe Lys Gly Gln Ile Lys Asn Leu Lys Ser Met Gly Glu Ser Trp Gly Leu Leu Ala Ser Asn Lys Ala Glu Val Phe Val Val Asn His Asp Arg Glu Arg Gly His Gly Gly Gly Gly Met Leu Thr Tyr Lys Asp Gly Ala Leu Tyr Asn Leu Ala Asn Ile Phe Met Leu Ala Trp Pro Tyr Gly Ala Tyr Pro Gln Val Met Ser Gly Tyr Asp Phe Gly Thr Asn Thr Asp Ile Gly Gly Pro Ser Ala Thr Pro Cys Ser Ser Gly Ser Ser Trp Asn Cys Glu His Arg Trp Ser Asn Ile Ala Asn Met Val Ser Phe His Asn Ala Ala Gln Gly Thr Ser Met Thr Asn Trp Trp Asp Asn Gly Asn Asn Gln Ile Ala Phe Gly Arg Gly Ala Lys Ala Phe Val Val Ile Asn Asn Glu Ser Ser Thr Leu Ser Lys Ser Leu Gln Thr Gly Leu Pro Ala Gly Glu Tyr Cys Asn Ile Leu Ala Gly Asp Ala Leu Cys Ser Gly Ser Thr Ile Lys Val Asp Ala Ser Gly Met Ala Thr Phe Asn Val Ala Gly Met Lys Ala Ala Ala Ile His Ile Asn Ala Lys Pro Asp Ser Thr Ser Ser Gly Ser Ser Gly Ser Ser Ser Gly Ser Ser Ser Ser Ala Thr Ser Asn Lys Phe Ala Ser Met Asn Leu Arg Gly Thr Asn Asn Gly Trp Ala Ser Thr Ala Met Thr Val Asp Ala Asn Arg Val Trp Ser Ala Asp Val Thr Phe Thr Gly Ala Ala Asp Ala Asn Gly Ala Gln Arg Phe Lys Phe Asp Val Tyr Gly Asn Trp Thr Glu Ser Tyr Gly Asp Thr Gln Ala Asp Gly Ile Ala Asp Lys Gly Ser Ala Lys Asp Ile Tyr Phe Asn Gly Val Gly Lys Tyr Arg Val Ser Leu Lys Glu Ser Asp Met Ser Tyr Thr Leu Thr Gln Leu Ser Ser Asn Gln Ala Pro Val Ala Ala Ile Thr Pro Lys Thr Leu Ser Val Lys Leu Gly Asp Ser Val Val Phe Asp Ala Ser Gly Ser Thr Asp Asp Val Gly Val Thr Gly Tyr Ser Trp Ser Thr Gly Gly Ser Ala Lys Thr Glu Thr Val Leu Phe Asp Ala Leu Gly Thr Lys Thr Ile Thr Val Thr Val Ala Asp Ala Asp Gly Leu Thr Ser Lys Ala Ser Ala Thr Val Thr Val Thr Asp Gly Ser Val Ala Tyr Asn Ser Asn Phe Ala Ser Leu Asn Phe Arg Gly Thr Pro Asn Ser Trp Gly Ala Ala Ala Met Thr Leu Val Ala Asp Asn Thr Trp Glu Ala Thr Val Asn Phe Asp Gly Gln Ala Asn Gln Arg Phe Lys Phe Asp Ile Lys Gly Asp Trp Ser Gln Asn Tyr Gly Asp Ser Asn Lys Asp Gly Val Ala Glu Arg Thr Gly Ala Asp Ile Tyr Thr Thr Val Thr Gly Gln Tyr Lys Val Gln Phe Asn Asp Ser Thr Leu Lys Tyr Thr Leu Thr Lys Leu Ala Asp Ser Ser Ala Thr Ser Tyr Ser Ala Asn Phe Ala Ser Leu Tyr Leu Arg Gly Thr Pro Asn Ser Trp Gly Thr Thr Ala Met Lys Leu Val Ala Asn Asn Ser Trp Gln Ala Glu Val Thr Phe Thr Gly Lys Gly Asp Ala Thr Gly Ala Gln Arg Phe Lys Phe Asp Val Lys Gly Asp Trp Ser Gln Asn Tyr Gly Asp Ser Asn Met Asp Gly Thr Ala Glu Arg Thr Gly Gly Asp Ile Thr Ser Ala Val Val Gly Thr Tyr Leu Val Thr Phe Asn Asp Ser Thr Leu Lys Tyr Thr Leu Thr Ala Lys

SEQ ID NO: 205

atgtaccgcgtaatacctattattttgattatgagtatgattgtagcttgtgagtctccaaagaaaaaaacaaccgaaaccgctcaaccttcaacaaa
tgccgaaaaaccctttgtttgggaggctgccaatgtatatttttgttaactgaccgttttaacaacggtaacccaaacaatgacatcaatttttaatag
gactaaagaatcaggaaaactccgcaattttatgggaggcgatatcaagggcatcacccaaaaaataaatgaggggtattttagtaaactaggc
gttaatgccatctggcttacccccggttgttgaacaaatacatggcagtgttgatgaaggtaccggcaatacctatgcctttcatggctattgggcca
aagattggagaaaactagaccctaatttggcacaaaagaagaacttgccgaactggtggcaaactgccatgcaaaaggcatcaggatactttta
gatgtggtaataaacaacaccggccccggttaaccgaacaagacccggtttggggagaagattgggtacgtacaggccccgcagtgtacctatga
taattacaccaataccaccagttgcacgctggtagccaatttacctgatatacttacagaaagtaatgaaaatgtggccttaccaacctttttgttaga
taaatggaaagccgaaggcagattagagcaagaactaaagaacttgacgatttttttcccgcacaggccacccacgcgcaccccgcttttac
attattaaatggcttaccgattacatccgagaatttggggtagatgggtttagggttgataccgtaaaacataccgaagaaacggtttgggccgag
ttgtatgatgaagccgtaattgcttttgccgaatataaaaaagccaacccagacaaggtattggacgataatgaatttatatggtaggcgaagtgt
acaactacggtatttccggcggaaggttctatgatttcggcgataaaaaggtggactattttgaccacggatttaaaagcctcatcaatttttgaaatg
aaatatgatgccaattttacctacgatacacttttaggaagtacgataccctttgcataccaaacttaaaggcagaagtgtgctcaactacctctca
tctcacgacgatggaagtccatttgataaaatgcggcaaaaaccatacgagtcggctacaaaattactgctcactccgggcgcatcccaaatttat
tacggtgacgaaaccgccagaagccttaacatagaaggcgcacagggagatgctacgcttcgttcgtttatgaattgggaagagctcgcagaa
gaccctgccaagcaaaaaatacttcagcattggcaaaaactgggcagtttcaggaacaaccaccccgcagttggtgccggaaggcacaaaac
ccttggcaaaaagccgttttacaccttagcagggtttatcaaaaaaatggttttattgacaaagttgtggtagcattagatgcccctaaaggccaaa
aacaaattaccgttaatggtgtttttgatgacggtacaaaacttgtagatgcctattcaggcaaagaaacctcagttaaaaatggtatcgtttcacttt
cttctgaatttgatattgtttttgttagaacaaaaataa

FIGURE 16KKKK

SEQ ID NO: 206

Met Tyr Arg Val Ile Pro Ile Ile Leu Ile Met Ser Met Ile Val Ala Cys Glu Ser Pro Lys Lys Lys Thr Thr Glu Thr Ala Gln Pro Ser Thr Asn Ala Glu Lys Pro Phe Val Trp Glu Ala Ala Asn Val Tyr Phe Leu Leu Thr Asp Arg Phe Asn Asn Gly Asn Pro Asn Asn Asp Ile Asn Phe Asn Arg Thr Lys Glu Ser Gly Lys Leu Arg Asn Phe Met Gly Gly Asp Ile Lys Gly Ile Thr Gln Lys Ile Asn Glu Gly Tyr Phe Ser Lys Leu Gly Val Asn Ala Ile Trp Leu Thr Pro Val Val Glu Gln Ile His Gly Ser Val Asp Glu Gly Thr Gly Asn Thr Tyr Ala Phe His Gly Tyr Trp Ala Lys Asp Trp Thr Asn Leu Asp Pro Asn Phe Gly Thr Lys Glu Asp Leu Ala Glu Leu Val Ala Thr Ala His Ala Lys Gly Ile Arg Ile Leu Leu Asp Val Val Ile Asn His Thr Gly Pro Val Thr Asp Gln Asp Pro Val Trp Gly Glu Asp Trp Val Arg Thr Gly Pro Gln Cys Thr Tyr Asp Asn Tyr Thr Asn Thr Thr Ser Cys Thr Leu Val Ala Asn Leu Pro Asp Ile Leu Thr Glu Ser Asn Glu Asn Val Ala Leu Pro Thr Phe Leu Leu Asp Lys Trp Lys Ala Glu Gly Arg Leu Glu Gln Glu Leu Lys Glu Leu Asp Asp Phe Phe Ser Arg Thr Gly His Pro Arg Ala Pro Arg Phe Tyr Ile Ile Lys Trp Leu Thr Asp Tyr Ile Arg Glu Phe Gly Val Asp Gly Phe Arg Val Asp Thr Val Lys His Thr Glu Glu Thr Val Trp Ala Glu Leu Tyr Asp Glu Ala Val Ile Ala Phe Ala Glu Tyr Lys Lys Ala Asn Pro Asp Lys Val Leu Asp Asp Asn Glu Phe Tyr Met Val Gly Glu Val Tyr Asn Tyr Gly Ile Ser Gly Gly Arg Phe Tyr Asp Phe Gly Asp Lys Lys Val Asp Tyr Phe Asp His Gly Phe Lys Ser Leu Ile Asn Phe Glu Met Lys Tyr Asp Ala Asn Phe Thr Tyr Asp Thr Leu Phe Arg Lys Tyr Asp Thr Leu Leu His Thr Lys Leu Lys Gly Arg Ser Val Le[illegible]n Tyr Leu Ser Ser His Asp Asp Gly Ser Pro Phe Asp Lys Met Arg Gln Lys Pro Tyr Glu Ser Ala [illegible]r Lys Leu Leu Leu Thr Pro Gly Ala Ser Gln Ile Tyr Tyr Gly Asp Glu Thr Ala Arg Ser Leu Asn Ile Glu Gly Ala Gln Gly Asp Ala Thr Leu Arg Ser Phe Met Asn Trp Glu Glu Leu Ala Glu Asp Pro Ala Lys Gln Lys Ile Leu Gln His Trp Gln Lys Leu Gly Ser Phe Arg Asn Asn His Pro Ala Val Gly Ala Gly Arg His Lys Thr Leu Gly Lys Lys Pro Phe Tyr Thr Phe Ser Arg Val Tyr Gln Lys Asn Gly Phe Ile Asp Lys Val Val Val Ala Leu Asp Ala Pro Lys Gly Gln Lys Gln Ile Thr Val Asn Gly Val Phe Asp Asp Gly Thr Lys Leu Val Asp Ala Tyr Ser Gly Lys Glu Thr Ser Val Lys Asn Gly Ile Val Ser Leu Ser Ser Glu Phe Asp Ile Val Leu Leu Glu Gln Lys.

SEQ ID NO: 207

ctgtcgactgagcctttcgttttgggctcgagactgactctcagcccaccccgcagtagctccagacggagtagccgtaatagccgttggccgg
gtcgtgggcaggggcctcgaggtacacccaccccgcttgagtccacccacttgtccacccagccgccgaggttgccggtgtactcgtggatgc
acgctcccgcgaacttcggaacgtagacccaccttccggctttgcttgaggcgaggttgatgtatgttatcagtcccggcttgcttccgtagccgt
ttctcacgaatatcagctcgtcgttgtcgtagtaaacgacgtcagtgcttcctccggccaggttgtcatgtatccagatgaggttcttgagcttatcct
tgttgagccactcctcgtagtcgcggtagaatattgtcggctggccctcgtaggtgaggatgaacgcgtaggctggatacttgttccagattatat
cggtgtcgtggtttgcaacgaaggttacggccttaaacgggtcgcggctgacgactgtgccccccgttcttgagggcctcgacgagtgcgggaa
tgttcttgttgtcaaaggccgcgtccatcttgtagtagagcgggaagtcgaagaccttggcgccgctcgagtaggcccagttgaggagtgcatc
aacgttggtgtcccagtactggcaacggccccaggcggcaacaccagttgagcaagtcctcgacgacccaggctccgtggccctcacgtagtc
aaagcgctaggcatcaacgccgatgctcctaggtaggcggcgtagctctcatcgctcgcccagagccagtgctggtcccagctcttctcgtg
ggctatgtctgggaagcctccaaatgtgccctcgtcacagcacttgacctcgttggggtggaagtcgaggtagttggcagtatatttgcccgagg
ccacctttgagaagtccgtccaggtgtagtccccaacgaacgggttccactcgaggtctccgcctgcgcggtggtttatgacgatgtccgctatg
acctttatgccgtaggcatgggccgtgtttatcatgttcacgagctcctgcttggagccaaagcgcgtctctaccgttcccttctggtcgtactcacc
gaggtcaaagaagtcgtaggggtcgtagcccatcgaataggcgccgcccatgcccttgctcgccgggggaatccaaatggcggatattcccg
cctcgtaccactccggtatcttgctcctgatggtgtcccaccagattcctccacctgggacgtcccagtagaaggcctgcattataacgccgccct
cttccagctcggagtacttggccataagttacctcctactagtagattaaaa

SEQ ID NO: 208

Leu Ser Thr Glu Pro Phe Val Leu Gly Ser Arg Leu Thr Leu Ser Pro Pro Arg Ser Ser Ser Arg Arg Ser Ser Arg Asn Ser Arg Trp Pro Gly Arg Gly Gln Gly Pro Arg Gly Thr Pro Thr Arg Leu Ser Pro Pro Thr Cys Pro Pro Ser Arg Arg Gly Cys Arg Cys Thr Arg Gly Cys Thr Leu Pro Arg Thr Ser Glu Arg Arg Pro Thr Phe Arg Leu Cys Leu Arg Arg Gly Cys Met Leu Ser Val Pro Ala Cys Phe Arg

FIGURE 16LLLL

Ser Arg Phe Ser Arg Ile Ser Ala Arg Arg Cys Arg Ser Lys Arg Arg Gln Cys Phe Leu Arg Pro Gly Cys His Val Ser Arg Gly Ser Ala Tyr Pro Cys Ala Thr Pro Arg Ser Arg Gly Arg Ile Leu Ser Ala Gly Pro Arg Arg Gly Thr Arg Arg Leu Asp Thr Cys Ser Arg Leu Tyr Arg Cys Arg Gly Leu Gln Arg Arg Leu Arg Pro Thr Gly Arg Gly Arg Leu Cys Pro Arg Ser Gly Pro Arg Arg Val Arg Glu Cys Ser Cys Cys Gln Arg Pro Arg Pro Ser Cys Ser Arg Ala Gly Ser Arg Arg Pro Trp Arg Arg Ser Ser Arg Pro Ser Gly Val His Gln Arg Trp Cys Pro Ser Thr Arg Gln Arg Pro Ser Arg Pro Thr Ser Ala Ser Pro Arg Pro Thr Leu Arg Gly Pro Ser Arg Ser Gln Ser Ala Arg His Gln Arg Arg Cys Ser Leu Gly Arg Arg Arg Ser Ser His Arg Ser Pro Arg Ala Ser Ala Gly Pro Ser Ser Ser Arg Gly Leu Cys Leu Gly Ser Leu Gln Met Cys Pro Arg His Ser Thr Pro Arg Trp Gly Gly Ser Arg Gly Ser Trp Gln Tyr Ile Cys Pro Arg Pro Pro Leu Arg Ser Pro Ser Arg Cys Ser Pro Gln Arg Thr Gly Ser Thr Arg Gly Leu Arg Leu Arg Gly Gly Leu Arg Cys Pro Leu Pro Leu Cys Arg Arg His Gly Pro Cys Leu Ser Cys Ser Arg Ala Pro Ala Trp Ser Gln Ser Ala Ser Leu Pro Phe Pro Ser Gly Arg Thr His Arg Gly Gln Arg Ser Arg Arg Gly Arg Ser Pro Ser Asn Arg Arg Arg Pro Cys Pro Cys Ser Pro Gly Glu Ser Lys Trp Arg Ile Phe Pro Pro Arg Thr Thr Pro Val Ser Cys Ser Trp Cys Pro Thr Arg Phe Leu His Leu Gly Arg Pro Ser Arg Arg Pro Ala Leu Arg Arg Pro Leu Pro Ala Arg Ser Thr Trp Pro Val Thr Ser Tyr Ile Lys

SEQ ID NO: 209

atgattcagcccatgcactctcgggaacaggcctgccgtctcattccggcactgatcatgacatttgcactggcactgccgttgcaaattcgtgccg
atgtcaccctgcatgctttcaactggagctatgccgatgtcgctgatcgggccgttgacatcgctgcagcagggtacagtgccgtgctggtggccc
cgccacttcgatccgaaggcacggcctggtgggcgcgataccagccccaggatctccgccttatcgaccatccgctgggcaatacacatgacttc
gtcaacatgatcgatgctctcgatgatgtgggtgtgggcgtgtacgccgacatcgtgctcaaccacatggccaatgaggctgcacaaaggcctga
cctgaactacccctggtcaggcagtgcttgacgaatatgcttccgatcccggtcatttcgagggcttgaggctgttcggtaatctgagcttcaatttcct
gtcggaacatgatttcggacccgcccagtgcattcaggattacagcgatgtgtttcaggtccagaactggcggctgtgcggaccgccgccggacc
cgggcctgcccgacctggtcgccaatgactgggtgatctctcaacagcgccagtatctggaagccatcaaggcgctgggtgtggctggcatgcg
catcgacgcggtcaagcatatgcccatgagccatatcaatgccgttctcacccccgagatccggtcgggcttgcatgtgtttggcgaagtcatcacc
tccggtggggctggtgatacatcctacgaccgttttctggcccccttacctggcacaaagcgaccatggtgcctatgactttccattgtttgaaaccatt
cgccgtgctttcggcttcggtggcagcatgagtgaactggtcgatcctgctgcctacggtcaggccctgccaccggaccgcgccatcaccttcgtc
atcacgcacgatattccgaacaatgacggatttcgctaccagatactcgaccccgtcgatgaatcactggcctacgcctacattctgggccgcgatg
gcggtgtcccgcttctgtattccgacaacaatgaaagcggcgatggccgctggatcgatgcctggcaacgtccggatctggttgcaatggtcggct
tccacaatgcagtccacggtcaggacatggccgtgctttcacatgacgactgccacctgctgtttcggcgcggcagcctcgggattgtcggcatca
acaagtgcggccatgcactcagctcctgggtcaacatgaaccagagcgtactgtggtggtacgcggactacacagacgtgctcgacagcaacag
cgttgtcaacatccagtcatcctggcacgagttcatccttcccgcccgccaggcacgcctgtggttgcga

SEQ ID NO: 210

MIQPMHSREQACRLIPALIMTFALALPLQIRADVTLHAFNWSYADVADRAVDIAAAGYSA
VLVAPPLRSEGTAWWARYQPQDLRLIDPLGNETHDFVNMMDALDDVGVGVYADIVLNHM
ANEAAQRPDLNVPGQAVLDEYASDPGIFGLRLFGNLSFNPLSEHDFGPAQCIQDYSDVF
QVQNWRLCGPPPDPGLPDLVANDWVISQQRQYLEAIKALGVAGMRIDAVKHMPMSHINA
VLTPEIRSGLHVFGEVITSGGAGDTSYDRFLAPYLAQSDHGAYDFPLFETIRRAFGFGGSMS
ELVDPAAYGQALPPDRAITFVITHDIPNNDGFRYQILDPVDESLAYAYILGRDGGVPLLYSD
NNESGDGRWIDAWQRPDLVAMVGFHNAVHGQDMAVLSHDDCHLLFRRGSLGIVGINKC
GHALSSWVNMNQSVLWWYADYTDVLDSNSVVNIQSSWHEFILPARQARLWLR

SEQ ID NO: 211

GTGTTTCGTTCTGACACAGTTTCGCGTACCTGCATGTATGGTGCGCTGCGTAATGCCTA
CCAACCCGATCGGGTGTTTACTGGAGTCACGGTGCGGACATGCAACTTAAAAAAGCAT
GCTCATCGCCAGGCGCTGTTGTTCATCGTGACGCGGTGCCTGTGCCTGAAATCCAGGC
AGACCCATAAAAACAACAACAAACCGATAACAAACGACCCAAGCCTTCTAAGAGGAG
AAAACGGGATGGCTTTTAAACTACGCAAAAAGGCGCTCGTTGGCCTGTTCACGGCCGG

FIGURE
16MMMM

CGCAATGGTATATGCCGGTGCAGCGGCGAGTGGTGAAATCATTCTGCAGGGCTTCCAC
TGGCACTCCAAGTGGGGCGGCAACAATCAGGGTTGGTGGCAGGTGATGGAAGGTCAG
GCCAACACCATCGCCAACGCCGGCTTTACGCACGTGTGGTTCCCGCCGGTCCATAACT
CGGCCGATGCCGAGGGTTACCTACCCCGCGAGCTGAACAACCTCAACTCCAGCTATGG
CTCCGAAGCACAGCTGCGCAGCGCCATCCAGGCACTGAACAATCGCGGCGTGCATGCG
ATTGCCGATGTGGTCATGAACCACCGGGTGGGCTGCTCTGGCTGGGCGGATTTCTGTA
ACCCGGACTGGCCGACCTGGTACATCGTCGCCAATGATTCCTGGCCCGGTGGCCCGAA
AAGCCAGAACTGGGACACGGGTGAGACGTACCACGCCGCCCGTGACCTCGATCACGC
CAATCCGCAGGTGCGCAACGATATCTCGCACTACCTGAACAGCCGCCTCAAGGACGTC
GGCTTCTCCGGCTGGCGCTGGGACTATGCCAAGGGTTTCTGGCCCGGCTATGTCGGCG
AGTACAACTGGAACACCAACCCGAACTTCTGTGTGGGTGAGGTGTGGGACGATCTCGA
CCCCAACAATCCCAACCCGCACCGCCAGCAACTGGTGGACTGGGTTGATGCTACCGGT
GGCAGTTGTCACGTCTTCGACTTCACCACCAAGGGGCTGACGAACTATGCGCTGCAGC
ATGGCCAGTACTGGCGCCTGCAGGGTGATAATGGTGGCCCGGCTGGCGGCATCGGCTG
GTGGCCGCAACGCATGGTGACCTTCGTCGACAACCATGACACGGGCCCGAGCAATCAC
TGTGGTGACGGCCAGAACCTCTGGCCCGTGCCCTGTGACAAGGTCATGGAGGCGTATG
CCTACATCCTGACCCATCCGGGCGTGCCGTCGGTGTACTGGACGCACTTCTTCAACTGG
AATCTTGGTAGCGAGATCAGCCAGTTGATGCAGATCCGCAAGAACCAGGGCGTGCACT
CCGGTTCCGACGTCTGGATCGCCGAGGCCCGTCACGGCCTGTACGCCGCCTATATCAA
CGGTAATGTGGCGATGAAGATGGGCTGGGATAACTGGAGCCCGGGCTGGGGCTGGTC
GCTGGCGGCCTCCGGTAACAACTGGGCCGTCTGGACACGCTGA

SEQ ID NO: 212
VFRSDTVSRTCMYGALRNAYQPDRVFTGVTVRTCNLKKHAHRQALLFIVTRCLCLKSRQT
HKNNNKPITNDPSLLRGENGMAFKLRKKALVGLFTAGAMVYAGAAASGEIILQGFHWHS
KWGGNNQGWWQVMEGQANTIANAGFTHVWFPPVHNSADAEGYLPRELNNLNSSYGSEA
QLRSAIQALNNRGVHAIADVVMNHRVGCSGWADFCNPDWPTWYIVANDSWPGGPKSQN
WDTGETYHAARDLDHANPQVRNDISHYLNSRLKDVGFSGWRWDYAKGFWPGYVGEYN
WNTNPNFCVGEVWDDLDPNNPNPHRQQLVDWVDATGGSCHVFDFTTKGLTNYALQHGQ
YWRLQGDNGGPAGGIGWWPQRMVTFVDNHDTGPSNHCGDGQNLWPVPCDKVMEAYA
YILTHPGVPSVYWTHFFNWNLGSEISQLMQIRKNQGVHSGSDVWIAEARHGLYAAYINGN
VAMKMGWDNWSPGWGWSLAASGNNWAVWTR

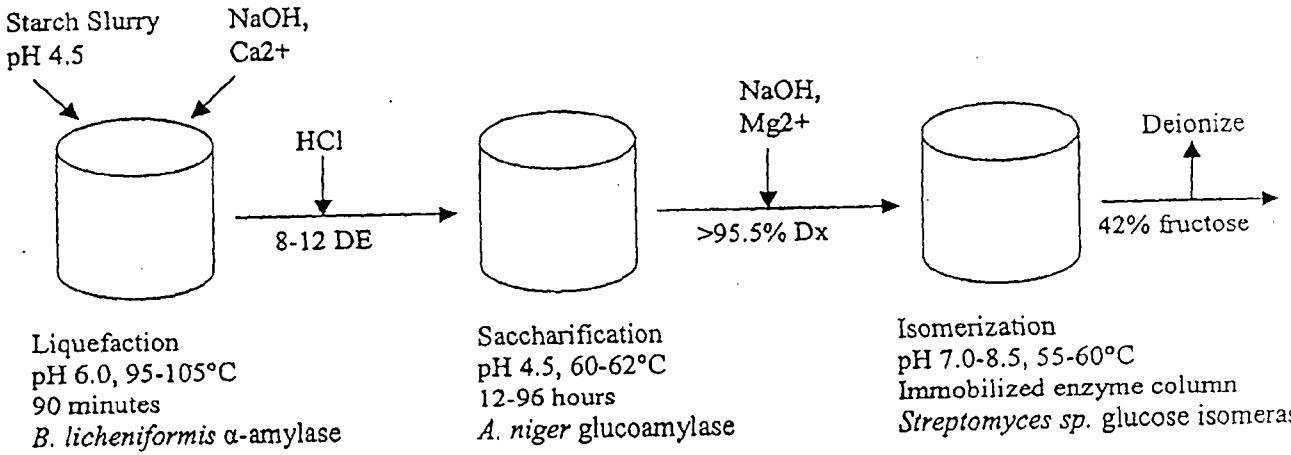

Starch Slurry
pH 4.5
NaOH,
Ca2+
HCl
8-12 DE
NaOH,
Mg2+
>95.5% Dx
Deionize
42% fructose
Liquefaction
pH 6.0, 95-105°C
90 minutes
B. licheniformis α-amylase
Saccharification
pH 4.5, 60-62°C
12-96 hours
A. niger glucoamylase
Isomerization
pH 7.0-8.5, 55-60°C
Immobilized enzyme column
Streptomyces sp. glucose isomerase

FIGURE 17

	SEQ ID NO.: 81	Pyro	Pyro	thermo	therm2	SEQ ID NO.: 75	SEQ ID NO.: 77	SEQ ID NO.: 83	SEQ ID NO.: 85	SEQ ID NO.: 79	SEQ 437
SEQ ID	100	91.7	75.1	82.1	80.1	82.5	82.6	82.1	82.6	83	77.8
pyro		100	74.8	82.5	80.5	82	82.2	82.9	82.8	84	78.5
Pyro2			100	71.5	71.1	74	74.2	77	77.1	73	70.5
therm				100	81.7	83.5	83.8	82.8	83.2	83.8	76.4
therm2					100	88.9	88.8	84.1	84.7	84	76.3
SEQ ID NO.: 75						100	98.3	84.6	85.2	85.5	77
SEQ ID							100	84.8	84.9	85.4	77.4
SEQ ID								100	96	83.3	78.5
SEQ ID									100	83	78.1
SEQ ID										100	79.8
Clone A											100

FIGURE 18

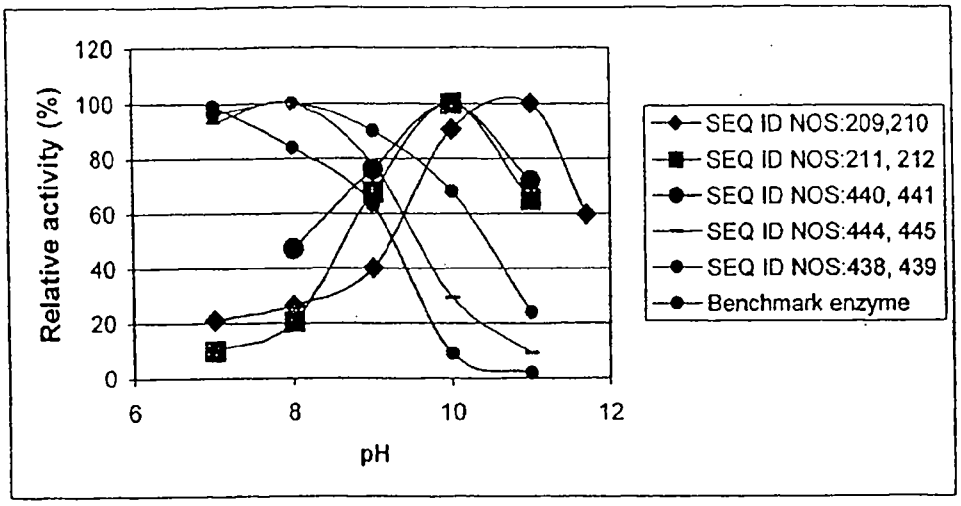

120
100
80
60
40
20
0
Relative activity (%)
6
8
10
12
pH
SEQ ID NOS:209,210
SEQ ID NOS:211, 212
SEQ ID NOS:440, 441
SEQ ID NOS:444, 445
SEQ ID NOS:438, 439
Benchmark enzyme

FIGURE 19

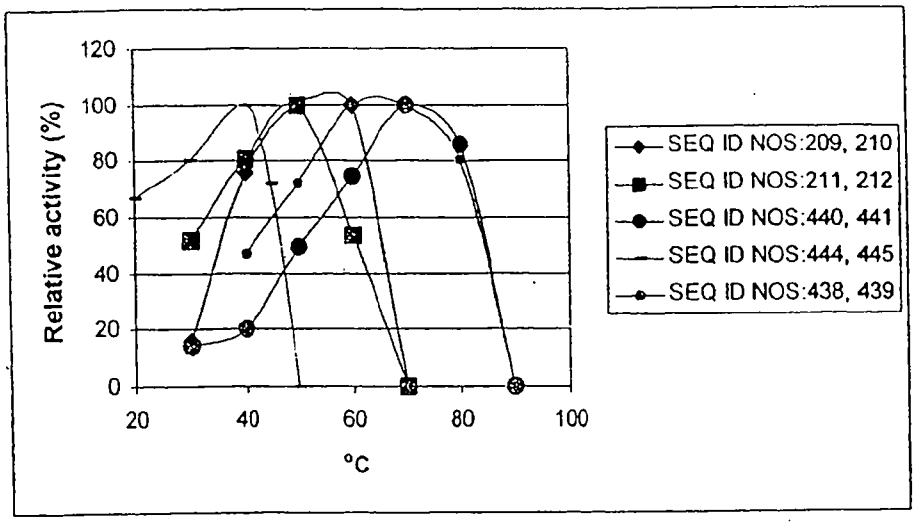

120
100
80
60
40
20
0
Relative activity (%)
20
40
60
80
100
°C
SEQ ID NOS:209, 210
SEQ ID NOS:211, 212
SEQ ID NOS:440, 441
SEQ ID NOS:444, 445
SEQ ID NOS:438, 439

FIGURE 20

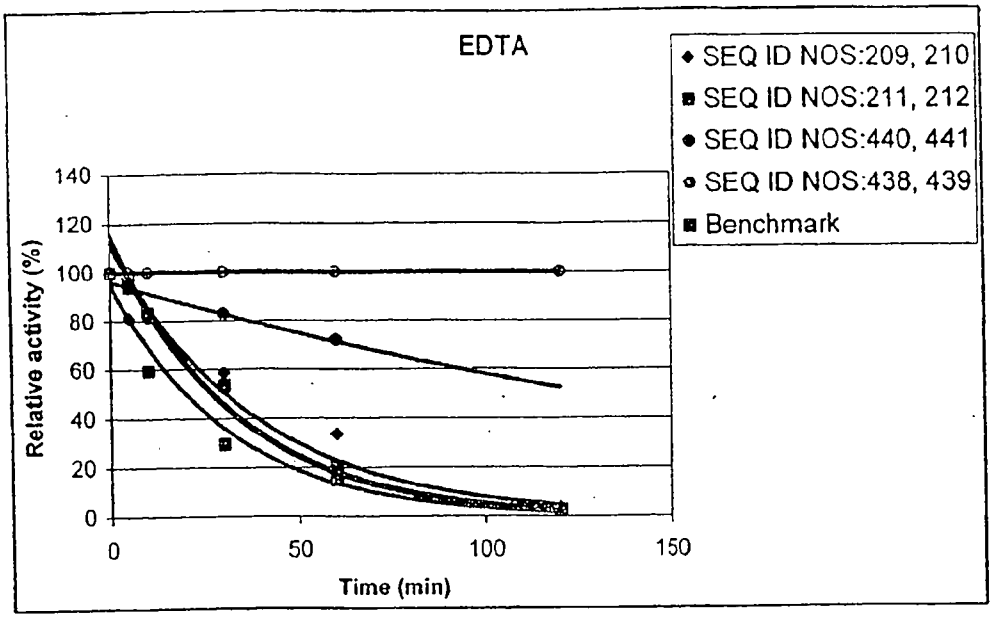

EDTA
SEQ ID NOS:209, 210
SEQ ID NOS:211, 212
SEQ ID NOS:440, 441
SEQ ID NOS:438, 439
Benchmark
Relative activity (%)
140
120
100
80
60
40
20
0
0
50
100
150
Time (min)

FIGURE 21

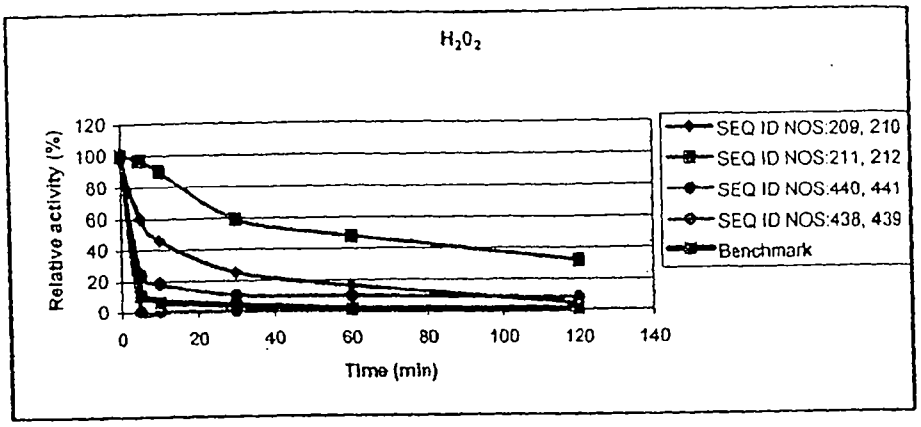

H_2O_2
Relative activity (%)
120
100
80
60
40
20
0
0
20
40
60
80
100
120
140
Time (min)
SEQ ID NOS:209, 210
SEQ ID NOS:211, 212
SEQ ID NOS:440, 441
SEQ ID NOS:438, 439
Benchmark

FIGURE 22

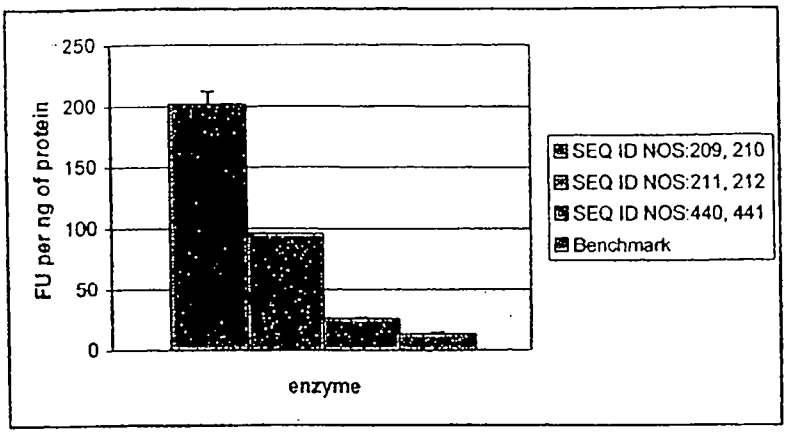

250
200
150
100
50
0
FU per ng of protein
enzyme
SEQ ID NOS:209, 210
SEQ ID NOS:211, 212
SEQ ID NOS:440, 441
Benchmark

FIGURE 23

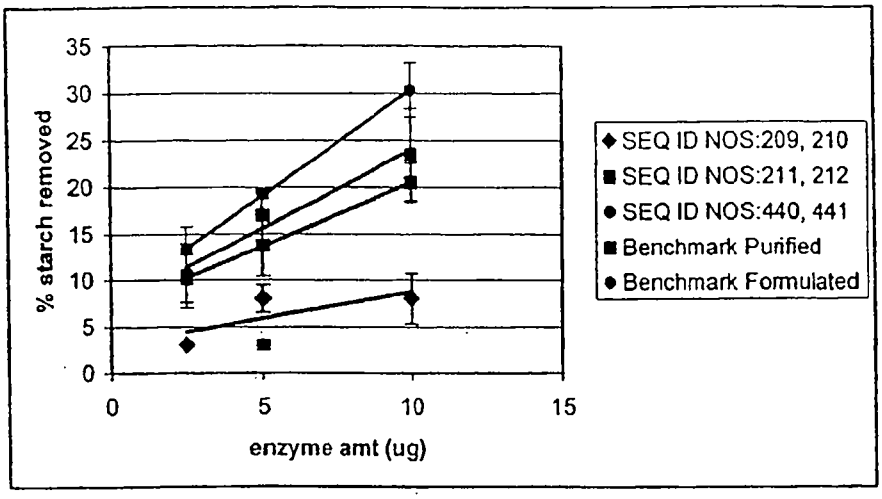

35
30
25
20
15
10
5
0
% starch removed
0
5
10
15
enzyme amt (ug)
SEQ ID NOS:209, 210
SEQ ID NOS:211, 212
SEQ ID NOS:440, 441
Benchmark Purified
Benchmark Formulated

FIGURE 24

FIGURE 25

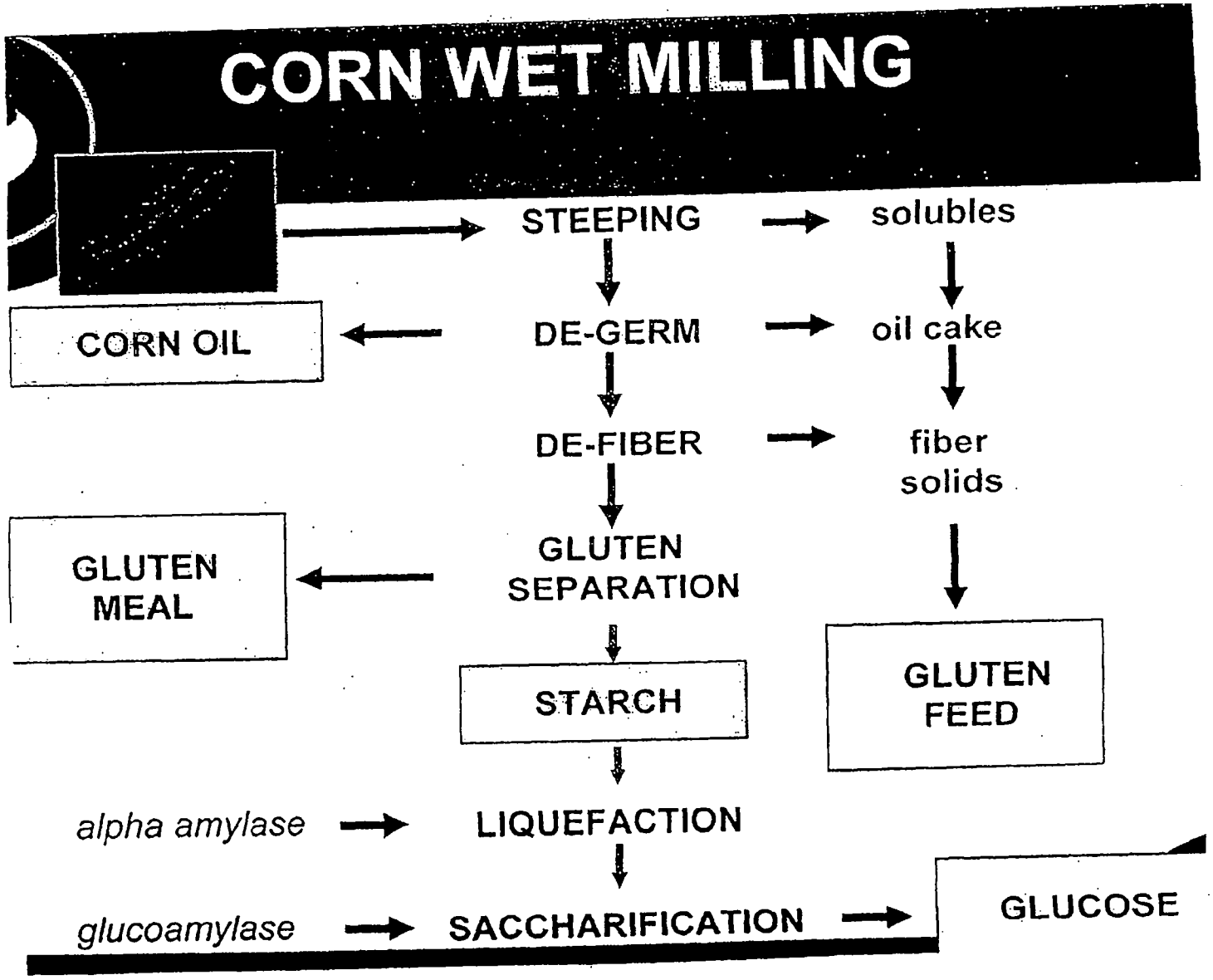

CORN WET MILLING
STEEPING
solubles
CORN OIL
DE-GERM
oil cake
DE-FIBER
fiber solids
GLUTEN MEAL
GLUTEN SEPARATION
STARCH
GLUTEN FEED
alpha amylase
LIQUEFACTION
glucoamylase
SACCHARIFICATION
GLUCOSE

FIGURE 26

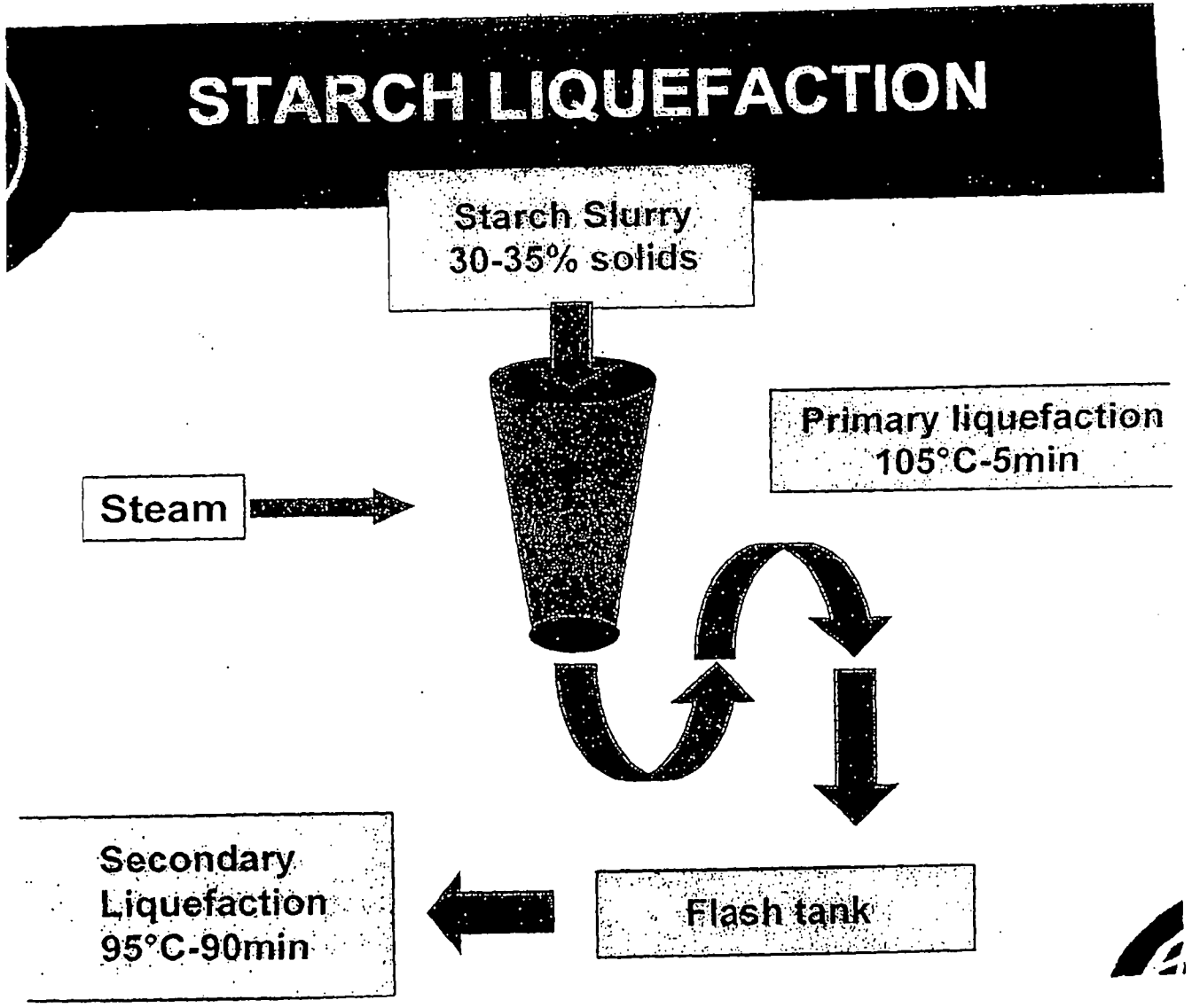

STARCH LIQUEFACTION
Starch Slurry
30-35% solids
Primary liquefaction
105°C-5min
Steam
Secondary
Liquefaction
95°C-90min
Flash tank

FIGURE 27

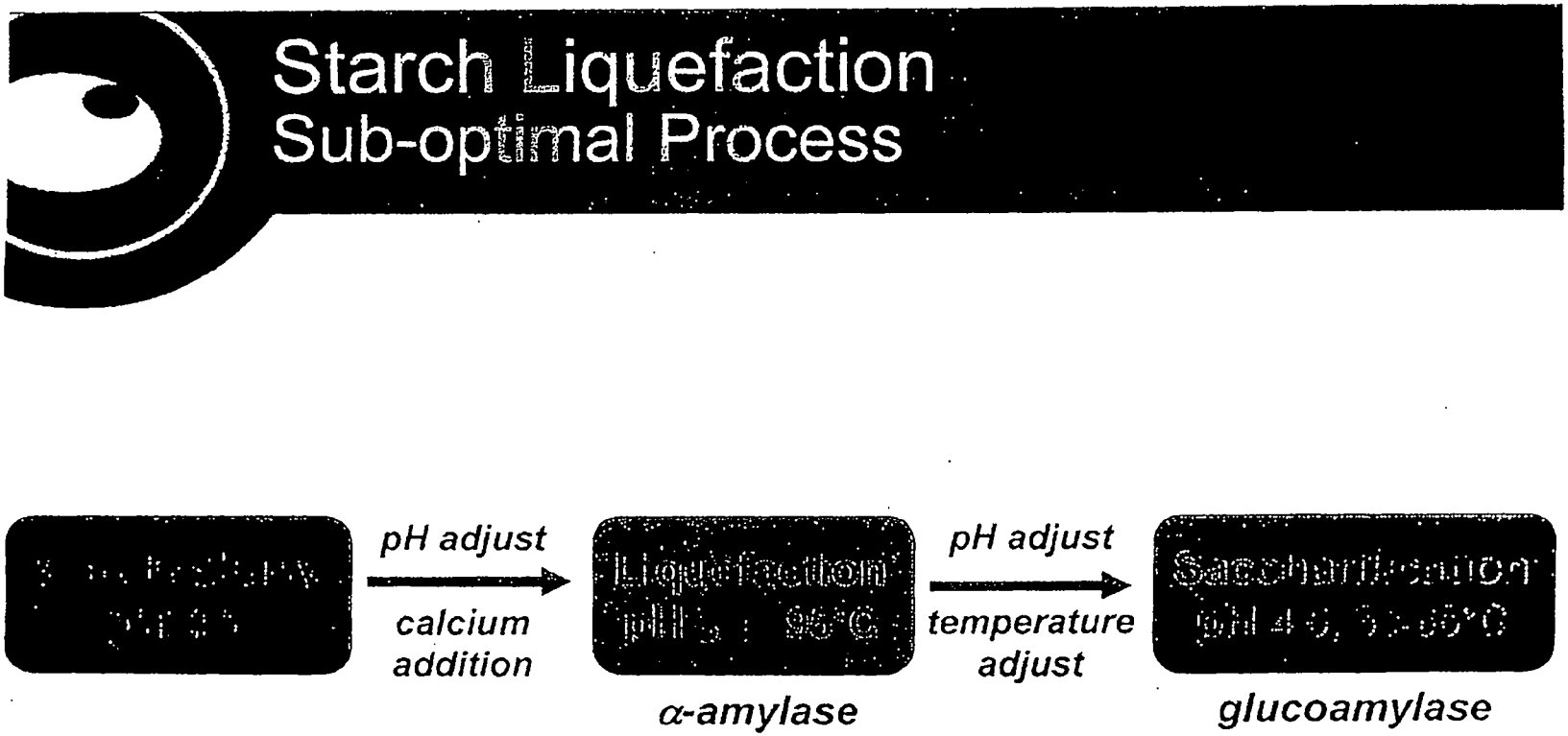
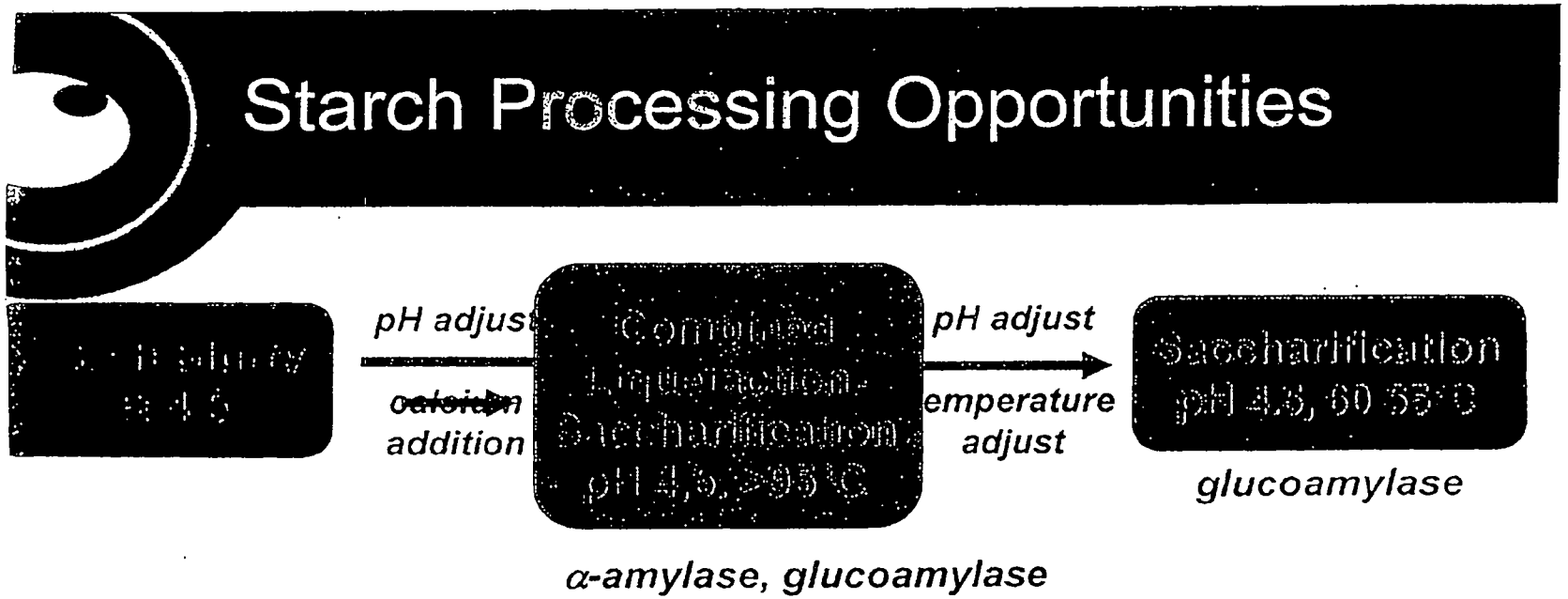

Starch Liquefaction
Sub-optimal Process
pH adjust
calcium addition
Liquefaction
pH 5.5 : 95°C
α-amylase
pH adjust
temperature adjust
Saccharification
pH 4.0, 55-60°C
glucoamylase

FIGURE 28


Starch Processing Opportunities
pH adjust
addition
pH adjust
adjust
Saccharification
glucoamylase
α-amylase, glucoamylase

Operating Conditions	D45	Termamyl® SC Amylase
Temperature	95°C	85°C
pH range	4.4 – 5.6	5.0 – 6.0 pH optimum 5.7
Ca++	No significant impact on viscosity reduction	Not required
Recycled Backset	Up to 30% well tolerated	Up to 30% well tolerated
Enzyme Dose	0.4 – 0.6 kg/MT starch	0.3 - 0.5 kg/MT starch
DE after liquefaction	6-10	12-14

FIGURE 29

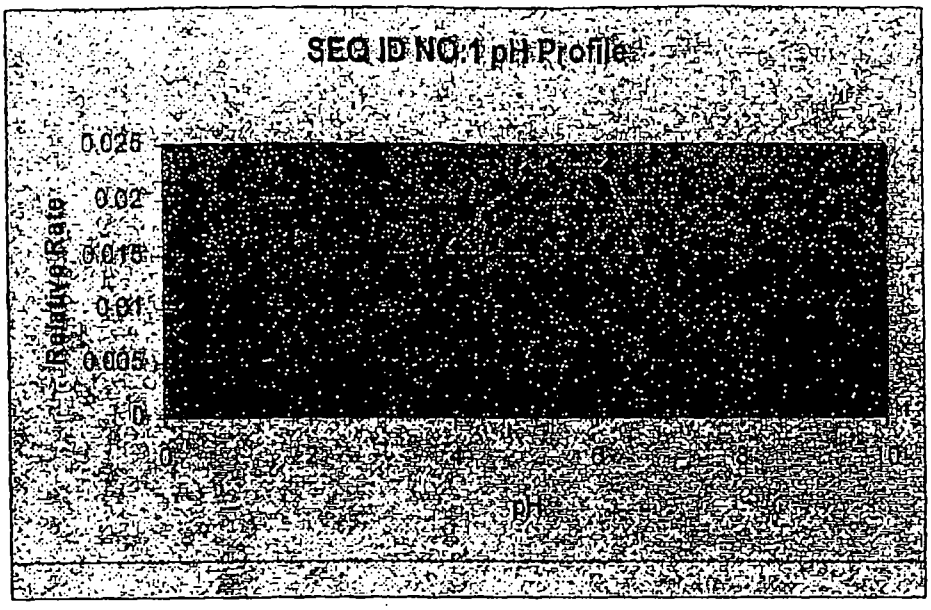

SEQ ID NO:1 pH Profile
Relative Rate
0.025
0.02
0.015
0.01
0.005
0
0
2
4
6
8
10
pH

FIGURE 30

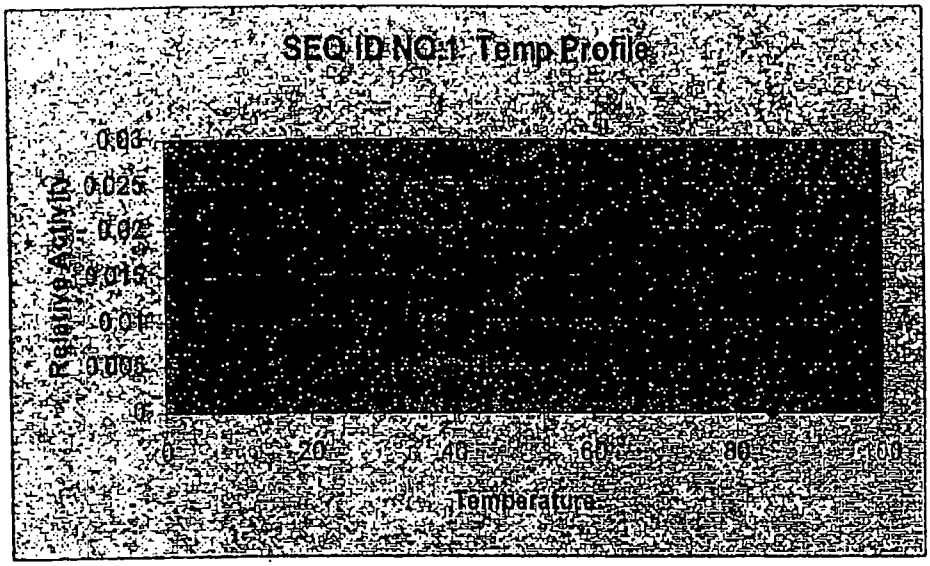

SEQ ID NO:1 Temp Profile
Relative Activity
0.03
0.025
0.02
0.015
0.01
0.005
0
0
20
40
60
80
100
Temperature

FIGURE 31

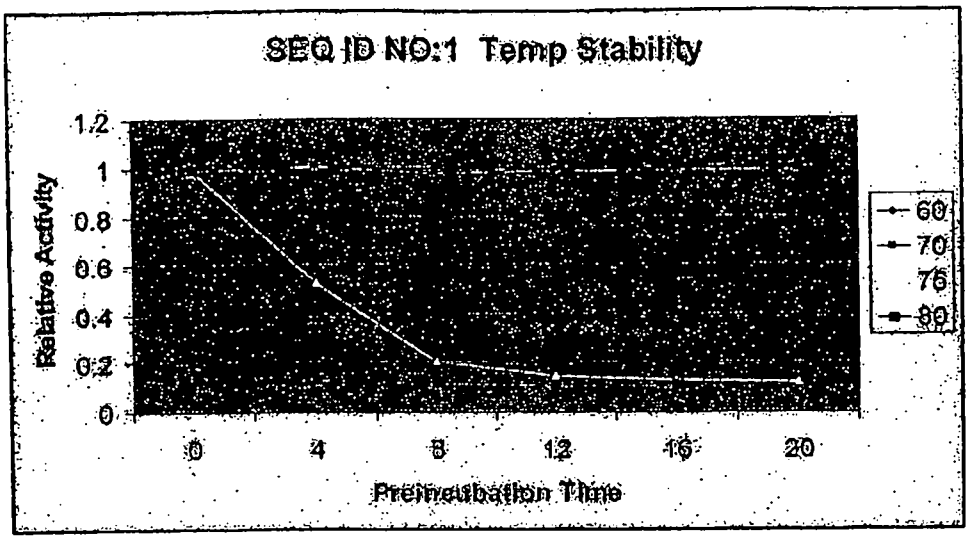

SEQ ID NO:1 Temp Stability
Relative Activity
1.2
1
0.8
0.6
0.4
0.2
0
0
4
8
12
16
20
Preincubation Time
60
70
75
80

FIGURE 32

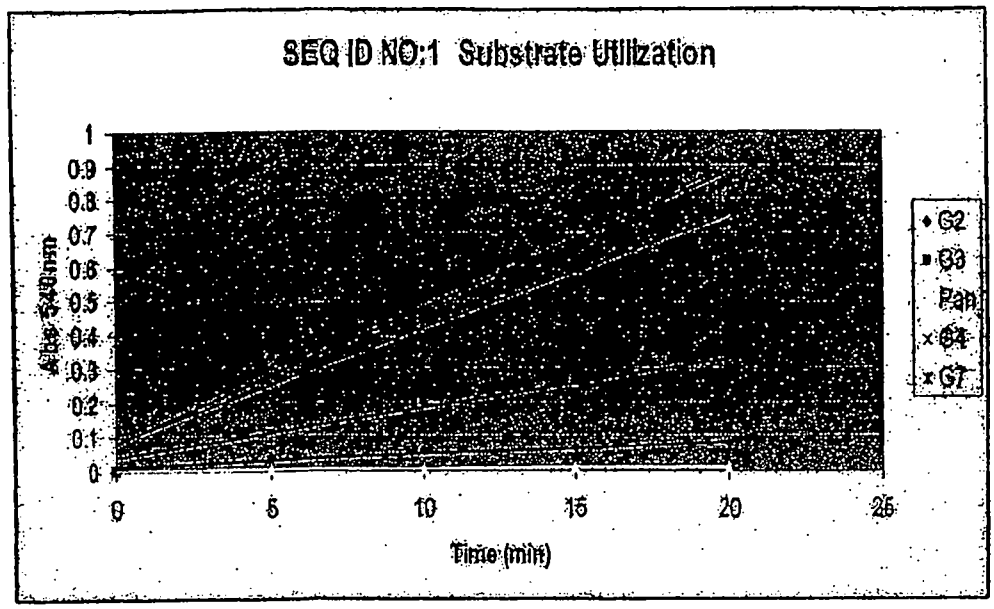

SEQ ID NO:1 Substrate Utilization
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
5
10
15
20
25
Time (min)
G2
G3
Pan
G4
G7

FIGURE 33

```
        *       *       30      *       *       60
tccaaagaattgaacgtcgtcggatgatcgtgatggtcagttttgaggtatcttccgaag

        *       *       90      *       *       120
agccaatttggattacccggcagtccggcagtccgacggtccagacaaggatcaatcagg

        *       *       150     *       *       180
gcctggcactattttagcatgtcggttagtccccaatcggctctgcatttgacagcggac

        *       *       210     *       *       240
atctgccgatcaggaacgaaacggtgcctgccgttgttgccctgtgagctctttatgcat

        *       *       270     *       *       300
gtcaattcataataatcccgataataggccgcagctattccaccctggggaaagggtttg

        *       *       330     *       *       360
ttgaatgtcggcgatttggggtacatctgggggttcggcacttcagtgtctccaggtccg

        *       *       390     *       *       420
aatttcatgcattcatgcatgaaaatgcggagggaacctcggcactgtcaccatccccgt

        *       *       450     *       *       480
caaactacggcttttggcgggctacattggctacaatggccacaatggctacaacggcct

        *       *       510     *       *       540
gtcaaggtatccattaatttctccgaagctcctgcaacgtttcgacaccctaaatttccg

        *       *       570     *       *       600
tttccgctacggagagcggagtaccggatacttttggacttttcatgtgcttctgtcgg

        *       *       630     *       *       660
gcattcaaagaccgggcctgtcctaacgtccaggggatcgtttggcagatcttccctcat

        *       *       690     *       *       720
```

FIGURE 34A

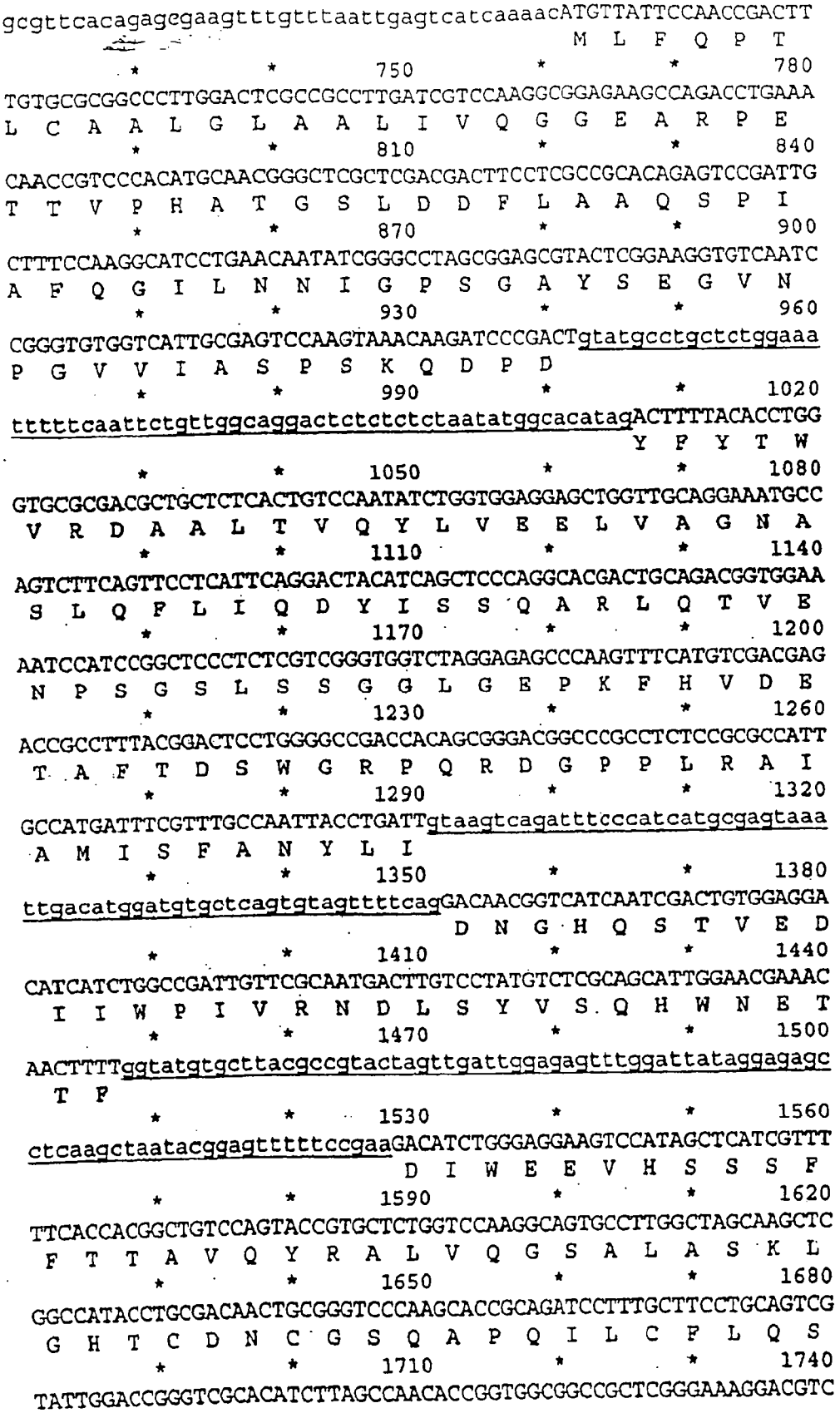


FIGURE 34B

```
 Y  W  T  G  S  H  I  L  A  N  T  G  G  G  R  S  G  K  D  V
                *           1770           *           *
AGCACGATCCTCGGCGTCATTGGCTCGTTTGATCCGAACGCCGACTGTGATGACGTTACC
 S  T  I  L  G  V  I  G  S  F  D  P  N  A  D  C  D  D  V  T
          *           *           1830           *           *
TTCCAGCCCTGCTCGGCCCGGGCTCTTGCAAATCACAAGCAGGTCGTTGACAGCTTCCGC
 F  Q  P  C  S  A  R  A  L  A  N  H  K  Q  V  V  D  S  F  R
          *           *           1890           *           *
AGTATCTATGCCATCAACGCTGGCATCCCGTCAGGGTCGGCTGTTGCGGTTGGACGTTAT
 S  I  Y  A  I  N  A  G  I  P  S  G  S  A  V  A  V  G  R  Y
          *           *           1950           *           *
CCCGAGGATGTCTATCAGGGTGGACACCCCTGGTACCTAACAACGGCTGCGGCGGCGGAG
 P  E  D  V  Y  Q  G  G  H  P  W  Y  L  T  T  A  A  A  A  E
          *           *           2010           *           *
CAGCTTTACGACGCCATTTACCAGTGGAACCATGTAGGGCACATCGACATCAATGCTGTC
 Q  L  Y  D  A  I  Y  Q  W  N  H  V  G  H  I  D  I  N  A  V
          *           *           2070           *           *
AATCTGGACTTCTTCAAGAGCATTTATCCGTCAGCCGCCGAGGGCACATACACATCAGAC
 N  L  D  F  F  K  S  I  Y  P  S  A  A  E  G  T  Y  T  S  D
          *           *           2130           *           *
TCTTCAACATTTCAAGACATTATATCTGCTGTACGGACCTATGCGGACGGGTTTCTCAGC
 S  S  T  F  Q  D  I  I  S  A  V  R  T  Y  A  D  G  F  L  S
          *           *           2190           *           *
GTAATTgtaagtccaaaccttcgaaaacgaatgcctcaagtcttccactgacattttgcg
 V  I
          *           *           2250           *           *
cagGAGAAATACACTCCGCCGGATAACTTGCTTGCCGAGCAGTTCCACCGGGAGACGGGC
    E  K  Y  T  P  P  D  N  L  L  A  E  Q  F  H  R  E  T  G
          *           *           2310           *           *
ATTCCACTATCGGCAGCTTCTCTGACATGGTCCTACGCCGCGCTCAACACGGCCGCGCAG
 I  P  L  S  A  A  S  L  T  W  S  Y  A  A  L  N  T  A  A  Q
          *           *           2370           *           *
CGGCGAGCGTCAATCGTGCCCTCACCGTGGAACTCTAACAGCACAGATCTCCCGGACAAA
 R  R  A  S  I  V  P  S  P  W  N  S  N  S  T  D  L  P  D  K
          *           *           2430           *           *
TGCTCGGCAACCTCGGCAACAGGGCCGTATGCCACGCCCACAAACACGGCATGGCCAACC
 C  S  A  T  S  A  T  G  P  Y  A  T  P  T  N  T  A  W  P  T
          *           *           2490           *           *
ACTACGCAGCCACCGGAGCGGCCGGCATGCACACCGCCGTCGGAAGTAACACTCACCTTC
 T  T  Q  P  P  E  R  P  A  C  T  P  P  S  E  V  T  L  T  F
          *           *           2550           *           *
AACGCGCTCGTCGACACCGCGTTTGGCCAGAATATTTATCTCGTGGGCTCCATTCCGGAG
 N  A  L  V  D  T  A  F  G  Q  N  I  Y  L  V  G  S  I  P  E
          *           *           2610           *           *
CTCGGATCGTGGGATCCGGCCAACGCCCTCTTGATGAGCGCAAAGAGCTGGACTAGCGGA
 L  G  S  W  D  P  A  N  A  L  L  M  S  A  K  S  W  T  S  G
          *           *           2670           *           *
AATCCGGTCTGGACGCTATCCATTTCCCTTCCAGCAGGAACCTCTTTTGAGTACAAGTTC
 N  P  V  W  T  L  S  I  S  L  P  A  G  T  S  F  E  Y  K  F
          *           *           2730           *           *
ATTCGAAAGGATGATGGTTCCTCGGATGTTGTCTGGGAAAGTGACCCGAATCGTTCGTAC
 I  R  K  D  D  G  S  S  D  V  V  W  E  S  D  P  N  R  S  Y
```

FIGURE 34C

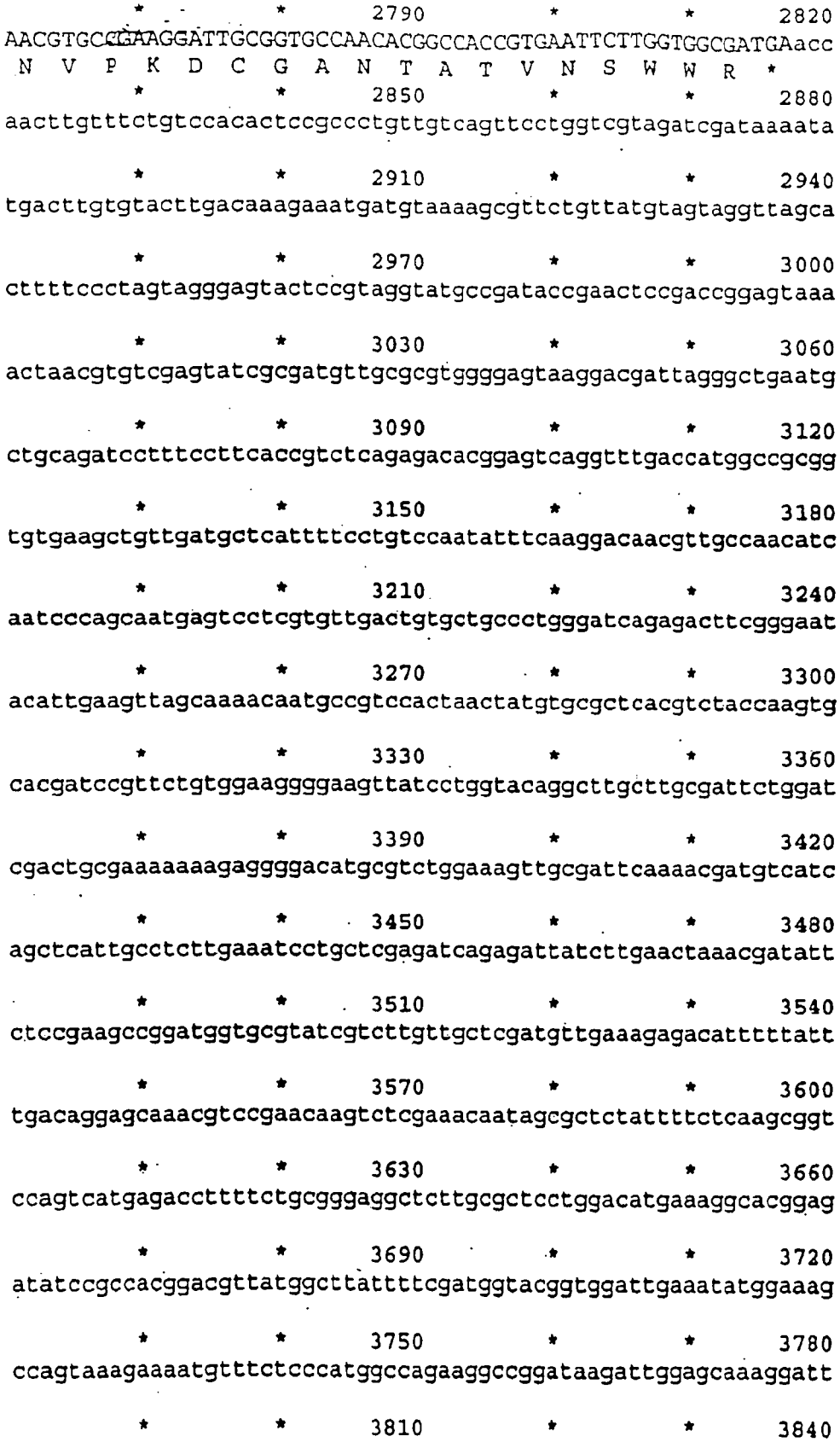

* * 2790 * * 2820
AACGTGCCGAAGGATTGCGGTGCCAACACGGCCACCGTGAATTCTTGGTGGCGATGAacc
N V P K D C G A N T A T V N S W W R *
* * 2850 * * 2880
aacttgtttctgtccacactccgccctgttgtcagttcctggtcgtagatcgataaaata
* * 2910 * * 2940
tgacttgtgtacttgacaaagaaatgatgtaaaagcgttctgttatgtagtaggttagca
* * 2970 * * 3000
cttttccctagtagggagtactccgtaggtatgccgataccgaactccgaccggagtaaa
* * 3030 * * 3060
actaacgtgtcgagtatcgcgatgttgcgcgtggggagtaaggacgattagggctgaatg
* * 3090 * * 3120
ctgcagatcctttccttcaccgtctcagagacacggagtcaggtttgaccatggccgcgg
* * 3150 * * 3180
tgtgaagctgttgatgctcattttcctgtccaatatttcaaggacaacgttgccaacatc
* * 3210 * * 3240
aatcccagcaatgagtcctcgtgttgactgtgctgccctgggatcagagacttcgggaat
* * 3270 * * 3300
acattgaagttagcaaaacaatgccgtccactaactatgtgcgctcacgtctaccaagtg
* * 3330 * * 3360
cacgatccgttctgtggaaggggaagttatcctggtacaggcttgcttgcgattctggat
* * 3390 * * 3420
cgactgcgaaaaaaagaggggacatgcgtctggaaagttgcgattcaaaacgatgtcatc
* * 3450 * * 3480
agctcattgcctcttgaaatcctgctcgagatcagagattatcttgaactaaacgatatt
* * 3510 * * 3540
ctccgaagccggatggtgcgtatcgtcttgttgctcgatgttgaaagagacatttttatt
* * 3570 * * 3600
tgacaggagcaaacgtccgaacaagtctcgaaacaatagcgctctattttctcaagcggt
* * 3630 * * 3660
ccagtcatgagacctttctgcgggaggctcttgcgctcctggacatgaaaggcacggag
* * 3690 * * 3720
atatccgccacggacgttatggcttattttcgatggtacggtggattgaaatatggaaag
* * 3750 * * 3780
ccagtaaagaaaatgtttctcccatggccagaaggccggataagattggagcaaaggatt
* * 3810 * * 3840

FIGURE 34D

gagatttattcgcggcggctatattatcatgaatggaagaagagagcagtcggaatgttg

* * 3870 * * 3900
gatctggaaactagaaagaggcggctttggcccgcaaacgataacacgtactttgaattc

* * 3930 * * 3960
tttgcatctgatcgatatcttttgatctcggggtaagttgaggacaggatagtagattcc

* * 3990 * * 4020
tgttggctcgacaggtcgtctaactgcagtgctaggaaacgtctgattgcgtgggatata

* * 4050 * * 4080
caggagtcgcggaaagttggggagattagtggtgagccctggcaggtgaacatcgtgcag

* * 4110
gataaggtcgtcatgtttgaacgaaagtcaa (SEQ ID NO:1)

FIGURE 34E

FIGURE 35A

69, 70	gi|12655802|gb|AAK00598.1|	0	alpha-amylase [Bacillus megaterium].	Bacillus megaterium	NO_HITS	91	80	513	533
87, 88	gi|3994289|gb|AAC87717.1|AR008691	0	Sequence 6 from patent US 5763460.	Unknown. Unclassified.	NO_HITS	97	N/A	472	549
89, 90	gi|12008232|gb|AAG44799.1|AF272860_2	0	amylopullulanase [Geobacillus stearothermophilus]	Geobacillus stearothermophilus Bacteria	NO_HITS	75	70	1100	2018
91, 92	gi|2842326|gb|AAB86961.1|	0	alpha amylase [Geobacillus stearothermophilus].	Geobacillus stearothermophilus Bacteria	NO_HITS	99	98	549	549
93, 94	gi|12008232|gb|AAG44799.1|AF272860_2	0	amylopullulanase [Geobacillus stearothermophilus].	Geobacillus stearothermophilus Bacteria	NO_HITS	74	70	515	2018
95, 96	gi|15644583|ref|NP_229636.1|	e-145	alpha-amylase [Thermotoga maritima].	Thermotoga maritima Bacteria	NO_HITS	52	58	507	556
97, 98	gi|15644583|ref|NP_229636.1|	e-145	alpha-amylase [Thermotoga maritima].	Thermotoga maritima Bacteria	NO_HITS	29	33	976	556
99, 100	gi|15644583|ref|NP_229636.1|	e-144	alpha-amylase [Thermotoga maritima].	Thermotoga maritima Bacteria	NO_HITS	53	58	511	556
101, 102	gi|13274566|gb|AAK17994.1|AF333075_1	0	[Pseudomonas sp. KFCC10818].	KFCC10818 Bacteria	NO_HITS	61	58	711	765
103, 104	gi|2842326|gb|AAB86961.1|	0	alpha amylase [Geobacillus stearothermophilus]	Geobacillus stearothermophilus Bacteria	NO_HITS	99	99	549	549
105, 106	gi|722279|gb|AAA63900.1|	0	alpha-amylase [Bacillus sp. TS-23].	Bacillus sp. TS-23 Bacteria	NO_HITS	84	76	550	613
107, 108	gi|1091118|prf||2020344A	3E-94	alpha amylase.	Thermoactinomyces vulgaris Bacteria	NO_HITS	37	N/A	502	453
109, 110	gi|4049918|gb|AAC97877.1|	0	[Thermococcus hydrothermalis].	hydrothermalis Archaea	NO_HITS	90	83	457	457
111, 112	gi|2251108|db||BAA21130.1|	0	alpha-amylase [Pyrococcus sp.].	Pyrococcus sp. Archaea	NO_HITS	83	77	471	461
113, 114	gi|113813|sp|P08278|AMY_BACLI	0	PRECURSOR (1,4-ALPHA-D-GLUCAN	licheniformis Bacteria	NO_HITS	99	99	512	512
115, 116	gi|2251108|db||BAA21130.1|	0	alpha-amylase [Pyrococcus sp.].	Pyrococcus sp. Archaea	NO_HITS	83	N/A	446	461
117, 118	gi|113810|sp|P22630|AMY1_AERHY	0	PRECURSOR (1,4-ALPHA-D-GLUCAN GLUCANOHYDROLASE).	Aeromonas hydrophila Bacteria	NO_HITS	84	81	491	484
119, 120	gi|722279|gb|AAA63900.1|	0	alpha-amylase [Bacillus sp. TS-23].	Bacillus sp. TS-23 Bacteria	NO_HITS	93	90	564	613
121, 122	gi|2253595|gb|AAC79122.1|	e-101	alpha-amylase [Drosophila ananassae].	ananassae Eukaryota	NO_HITS	42	63	617	494
123, 124	gi|113822|sp|P22998|AMY_STRVL	e-152	ALPHA-AMYLASE PRECURSOR (1,4-ALPHA-D-GLUCAN	Streptomyces violaceus Bacteria	NO_HITS	46	59	589	589
125, 126	gi|2251108|db||BAA21130.1|	0	alpha-amylase [Pyrococcus sp.].	Pyrococcus sp. Archaea	NO_HITS	84	78	484	461
127, 128	gi|722279|gb|AAA63900.1|	0	alpha-amylase [Bacillus sp. TS-23].	Bacillus sp. TS-23 Bacteria	NO_HITS	94	84	615	613

FIGURE 35B

129, 130	gi|7271927|gb|AAF44693.1|AF240464_1	5E-88	alpha-amylase [Pyrococcus woesei]	Pyrococcus woesei Archaea	NO_HITS	26	38	617	460
131, 132	gi|535792|dbj|BAA01600.1|	e-171	maltopentaose forming amylase [Pseudomonas sp.]	Pseudomonas sp. Bacteria	NO_HITS	61	76	626	614
133, 134	gi|722279|gb|AAA63900.1|	0	alpha-amylase [Bacillus sp. TS-23]	Bacillus sp. TS-23 Bacteria	NO_HITS	97	94	548	613
135, 136	gi|722279|gb|AAA63900.1|	0	alpha-amylase [Bacillus sp. TS-23]	Bacillus sp. TS-23 Bacteria	NO_HITS	84	84	544	613
137, 138	gi|207462|gb|AAA42272.1|	0.48	tropoelastin [Rattus norvegicus]	Rattus norvegicus Eukaryota	NO_HITS	19	11	439	307
139, 140	gi|11226329|emb|CAC16485.1|	e-173	[Geobacillus stearothermophilus]	stearothermophilus Bacteria	NO_HITS	60	59	507	515
141, 142	gi|15601016|ref|NP_232648.1|	0	alpha-amylase [Vibrio cholerae]	Vibrio cholerae Bacteria	NO_HITS	70	67	466	486
143, 144	gi|15601016|ref|NP_232648.1|	0	alpha-amylase [Vibrio cholerae]	Vibrio cholerae Bacteria	NO_HITS	69	65	473	486
145, 146	gi|12655802|gb|AAK00698.1|	0	alpha-amylase [Bacillus megaterium]	Bacillus megaterium Bacteria	NO_HITS	91	90	513	533
147, 148	gi|9081816|gb|AAF82611.1|	1E-77	beta-agarase [Pseudomonas sp. W7]	Pseudomonas sp. W7 Bacteria	NO_HITS	28	39	781	642
149, 150	gi|9081816|gb|AAF82611.1|	3E-72	beta-agarase [Pseudomonas sp. W7]	Pseudomonas sp. W7 Bacteria	NO_HITS	24	36	834	642
151, 152	gi|14597796|emb|CAC43721.1|	4E-32	c_pp010010057r [Physcomitrella patens]	Physcomitrella patens Eukaryota	NO_HITS	16	28	704	430
153, 154	gi|113814|sp|P20845|AMY_BACME	0	ALPHA-AMYLASE PRECURSOR (1,4-ALPHA-D-GLUCAN	Bacillus megaterium Bacteria	NO_HITS	76	73	539	520
155, 156	gi|6502675|gb|AAF14358.1|	0	alpha-amylase [Bacillus subtilis]	Bacillus subtilis Bacteria	NO_HITS	97	96	591	659
157, 158	gi|113814|sp|P20845|AMY_BACME	0	PRECURSOR (1,4-ALPHA-D-GLUCAN	megaterium Bacteria	NO_HITS	95	91	535	520
159, 160	gi|113814|sp|P20845|AMY_BACME	0	PRECURSOR (1,4-ALPHA-D-GLUCAN GLUCANOHYDROLASE)	Bacillus megaterium Bacteria	NO_HITS	96	93	526	520
161, 162	gi|113814|sp|P20845|AMY_BACME	0	ALPHA-AMYLASE PRECURSOR (1,4-ALPHA-D-GLUCAN	Bacillus megaterium Bacteria	NO_HITS	95	98	540	520
163, 164	gi|12655802|gb|AAK00698.1|	0	alpha-amylase [Bacillus megaterium]	Bacillus megaterium Bacteria	NO_HITS	69	64	473	533
165, 166	gi|12655802|gb|AAK00698.1|	0	alpha-amylase [Bacillus megaterium]	Bacillus megaterium Bacteria	NO_HITS	89	88	476	533
167, 168	gi|113814|sp|P20845|AMY_BACME	0	PRECURSOR (1,4-ALPHA-D-GLUCAN GLUCANOHYDROLASE)	Bacillus megaterium Bacteria	NO_HITS	96	92	531	520
169, 170	gi|11226329|emb|CAC16485.1|	e-166	unnamed protein product [Geobacillus stearothermophilus]	Geobacillus stearothermophilus Bacteria	NO_HITS	60	59	506	515
171, 172	gi|113814|sp|P20845|AMY_BACME	0	ALPHA-AMYLASE PRECURSOR (1,4-	Bacillus megaterium	NO_HITS	98	94	476	520
173, 174	gi|113814|sp|P20845|AMY_BACME	0	ALPHA-AMYLASE PRECURSOR (1,4-ALPHA-D-GLUCAN GLUCANOHYDROLASE)	Bacillus megaterium Bacteria	NO_HITS	96	93	531	520

FIGURE 35C

175, 176	gi|15601018|ref|NP_232648.1|	0	alpha-amylase [Vibrio cholerae]	Vibrio cholerae Bacteria	NO_HITS	73	68	465	466
177, 178	gi|11228329|emb|CAC16485.1|	e-187	unnamed protein product [Geobacillus stearothermophilus]	Geobacillus stearothermophilus Bacteria	NO_HITS	60	59	507	515
179, 180	gi|17905898|gb|AAE61102.1|	e-166	Sequence 6 from patent US 6297826.	Unknown. Unclassified.	NO_HITS	59	N/A	507	514
181, 182	gi|113760|sp|P22963|AMT4_PSESA	0	GLUCAN 1,4-ALPHA-MALTOTETRAHYDROLASE PRECURSOR (G4-	Pseudomonas saccharophila Bacteria	NO_HITS	89	N/A	809	651
183, 184	gi|113814|sp|P20845|AMY_BACME	0	ALPHA-AMYLASE PRECURSOR (1,4-ALPHA-D-GLUCAN GLUCANOHYDROLASE).	Bacillus megaterium Bacteria	NO_HITS	95	92	531	520
185, 186	gi|113814|sp|P20845|AMY_BACME	0	PRECURSOR (1,4-ALPHA-D-GLUCAN GLUCANOHYDROLASE).	Bacillus megaterium Bacteria	NO_HITS	97	94	523	520
187, 188	gi|15226783|ref|NP_181616.1|	3E-66	4-alpha-glucanotransferase [Arabidopsis thaliana]	Arabidopsis thaliana Eukaryota	NO_HITS	29	28	683	738
189, 190	gi|113814|sp|P20845|AMY_BACME	0	ALPHA-AMYLASE PRECURSOR (1,4-ALPHA-D-GLUCAN GLUCANOHYDROLASE).	Bacillus megaterium Bacteria	NO_HITS	95	92	531	520
191, 192	gi|113814|sp|P20845|AMY_BACME	0	ALPHA-AMYLASE PRECURSOR (1,4-	Bacillus megaterium	NO_HITS	96	92	531	520
193, 194	gi|13274586|gb|AAK17994.1|AF333075_1	0	alpha-amylase 4 [Pseudomonas sp. KFCC10818]	Pseudomonas sp. KFCC10818 Bacteria	NO_HITS	77	66	653	765
195, 196	gi|482670|pir|A60999	0	3.2.1.1) precursor - Micrococcus sp. (strain	Micrococcus sp. Bacteria	3.2.1.1	59	56	929	1104
197, 198	gi|15601018|ref|NP_232648.1|	0	alpha-amylase [Vibrio cholerae]	Vibrio cholerae Bacteria	NO_HITS	72	69	486	486
199, 200	gi|14029135|gb|AAK51132.1|	1.6	orphan seven transmembrane receptor [Rattus norvegicus]	Rattus norvegicus Eukaryota	NO_HITS	27	44	132	730
201, 202	gi|13182951|gb|AAK15003.1|AF233372_1	1E-33	neopullulanase [Bacillus stearothermophilus]	Bacillus stearothermophilus Bacteria	NO_HITS	12	40	630	588
203, 204	gi|726846|sp|P41131|AMYA_AERHY	e-148	PRECURSOR (1,4-ALPHA-D-GLUCAN GLUCANOHYDROLASE).	Aeromonas hydrophila Bacteria	NO_HITS	32	35	866	443
205, 206	gi|97758|pir|S10789	e-101	amylase A-180 - alkaliphilic eubacterium 163-26.	alkaliphilic eubacterium 163-26 Bacteria	NO_HITS	42	54	567	1684

FIGURE 35D

207, 208	gi|15218510|ref|NP_174202.1|	0.94	plastid ribosomal protein L34 precursor, putative [Arabidopsis thaliana]	Arabidopsis thaliana Eukaryota	NO_HITS	11	10	439	157
209, 210	gi|477015|pir||A47574 a	e-175	alpha-amylase [Xanthomonas campestris]	Xanthomonas campestris	NO_HITS	61	58	472	475
211, 212	gi|15222859|ref|NP_177	2.10E-59	alpha-amylase, putative [Arabidopsis thaliana]	Arabidopsis thaliana Eukaryota	NO_HITS	42	38	498	413

FIGURE 35E

[illegible]	Sequence Top Hit	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
322, 323	alpha-amylase precursor [Thermococcus sp. GU5L5]	Thermococcus sp. GU5L5	3.2.1.1	21326995	0	2251107	1392	463	461	83
324, 325	hypothetical protein [Burkholderia fungorum]	Burkholderia fungorum	3.2.1.1	22986674	1E-28	7379424	1296	431	1146	25
326, 327	alpha-amylase (EC 3.2.1.1) precursor - Bacillus megaterium	Bacillus megaterium	3.2.1.1	80110	7E-98	11344494	1359	452	520	42
328, 329	hypothetical protein [Chloroflexus aurantiacus]	Chloroflexus aurantiacus	3.2.1.1	22970588	1E-155	535791	1677	558	575	53
330, 331	hypothetical protein [Chloroflexus aurantiacus]	Chloroflexus aurantiacus	2.4.1.-	22971468	0	4633806	3129	1042	851	55
332, 333	hypothetical protein [Microbulbifer degradans 2-40]	Microbulbifer degradans 2-40	3.2.1.1	23027235	1E-117	166984	1707	569	643	49
334, 335	hypothetical protein [Microbulbifer degradans 2-40]	Microbulbifer degradans 2-40	3.2.1.1	23027235	7E-91	62191	2061	686	643	43

FIGURE 35F

336, 337	alpha-amylase (EC 3.2.1.1) precursor - Streptomyces violaceus	Streptomyces violaceus	3.2.1.1	80864	0	153156	1731	576	569	65
338, 339	unnamed protein product	Bacillus sp.	3.2.1.1	13539158	0	14774986	1704	568	587	90
340, 341	outer membrane protein [Bacteroides thetaiotaomicron]	Bacteroides thetalotaomicron	3.2.1.10	1478030	1E-137	153158	1848	615	692	46
342, 343	hypothetical protein [Microbulbifer degradans 2-40]	Microbulbifer degradans 2-40	3.2.1.1	23027631	0	13274585	2061	686	563	64
344, 345	hypothetical protein [Microbulbifer degradans 2-40]	Microbulbifer degradans 2-40	3.2.1.1	23027631	1E-179	8247214	1980	659	563	62
346, 347	amylase precursor	Aeromonas hydrophila	3.2.1.1	141870	0	141869	1398	465	464	92
348, 349	alpha-amylase [Nostoc sp. PCC 7120]	Nostoc sp. PCC 7120	3.2.1.1	17229682	0	450848	1488	495	492	61
350, 351	ALPHA-AMYLASE PRECURSOR (1,4-ALPHA-D-GLUCAN GLUCANOHYDROLASE)	Pseudoalteromonas haloplanktis	3.2.1.1	6226551	0	2467084	2001	666	669	67
352, 353	hypothetical protein [Microbulbifer degradans 2-40]	Microbulbifer degradans 2-40	3.2.1.1	23027235	1E-135	3549647	1263	420	643	55

FIGURE 35G

354, 355	ALPHA-AMYLASE PRECURSOR (1,4-ALPHA-D-GLUCAN GLUCANOHYDRO LASE)	Aeromonas hydrophila	3.2.1.1	728848	0	304014	2577	858	443	44
356, 357	hypothetical protein [Microbulbifer degradans 2-40]	Microbulbifer degradans 2-40	3.2.1.1	23027235	7E-95	5442101	4875	1625	643	47
358, 359	Glycosidase [Vibrio vulnificus CMCP6]	Vibrio vulnificus	3.2.1.1	27366839	1E-147	155351	1422	473	466	57
360, 361	amylase	Bacillus thuringiensis	2.4.1.18	580662	0	2635411	1938	645	648	98
362, 363	hypothetical protein [Microbulbifer degradans 2-40]	Microbulbifer degradans 2-40	3.2.1.1	23027235	1E-65	13362592	2094	697	643	33
364, 365	Sequence 6 from patent US 5753460		3.2.1.1	3994289	0	722278	1536	511	549	69
366, 367	ALPHA-AMYLASE PRECURSOR (1,4-ALPHA-D-GLUCAN GLUCANOHYDRO LASE)	Pseudoalteromonas haloplanktis	3.2.1.1	6226551	0	2487084	1992	663	669	70
368, 369	hypothetical protein [Burkholderia fungorum]	Burkholderia fungorum	3.2.1.1	22986674	7E-28	14547281	1257	418	1146	27

FIGURE 35H

370, 371	alpha-amylase A [Halothermothrix orenii]	Halothermothrix orenii		24306106	1E-107	216309	1614	537	515	41
372, 373	hypothetical protein [Microbulbifer degradans 2-40]	Microbulbifer degradans 2-40	3.2.1.1	23027235	6E-77	166984	1437	478	643	42
374, 375	amylase precursor	Aeromonas hydrophila	3.2.1.1	141870	0	141869	1398	465	464	91
376, 377	hypothetical protein [Microbulbifer degradans 2-40]	Microbulbifer degradans 2-40	3.2.1.1	23027235	6E-52	13702782	1551	516	643	32
378, 379	hypothetical protein [Microbulbifer degradans 2-40]	Microbulbifer degradans 2-40	3.2.1.1	23027235	1E-124	20334	1269	422	643	53
380, 381	periplasmic alpha-amylase precursor [Xanthomonas campestris]	Xanthomonas campestris	3.2.1.1	1166403	1E-140	1166402	1644	547	526	49
382, 383	putative bi-functional protein (secreted alpha-amylase/dextrinase) [Streptomyces coelicolor A3(2)]	Streptomyces coelicolor A3(2)	3.2.1.1	21220698	0	288182	4176	1391	1798	50
384, 385	alpha-amylase [Xanthomonas campestris pv. campestris]	Xanthomonas campestris pv. campestris	3.2.1.1	19224331	0	155351	1434	477	475	65
386, 387	hypothetical protein [Chloroflexus aurantiacus]	Chloroflexus aurantiacus	3.2.1.1	22970588	1E-162	1771460	1458	485	575	59

FIGURE 35I

388, 389	GLUCAN 1,4-ALPHA-MALTOTETRAHYDROLASE PRECURSOR (G4-AMYLASE) (MALTOTETRAOSE-FORMING AMYLASE) (EXO-MALTOTETRAOHYDROLASE) (MALTOTETRAOSE-FORMING EXO-AMYLASE)	Pseudomonas stutzeri	3.2.1.1	2506188	0	45821	1662	553	548	92
390, 391	alpha-amylase [Xanthomonas axonopodis pv. citri str. 306]	Xanthomonas axonopodis pv. citri str. 306	3.2.1.1	21106921	1E-177	155351	1497	498	475	61
392, 393	hypothetical protein [Nostoc punctiforme]	Nostoc punctiforme	3.2.1.1	23126762	8E-28	7799230	2100	699	552	29
394, 395	hypothetical protein [Microbulbifer degradans 2-40]	Microbulbifer degradans 2-40	3.2.1.1	23027235	1E-147	14023709	1347	448	643	60
396, 397	hypothetical protein [Microbulbifer degradans 2-40]	Microbulbifer degradans 2-40	3.2.1.1	23027244	1E-151	11433676	1644	547	566	48
398, 399	hypothetical protein [Nostoc punctiforme]	Nostoc punctiforme	2.4.1.18	23126762	5E-28	13276803	2040	679	552	32
400, 401	hypothetical protein [Microbulbifer degradans 2-40]	Microbulbifer degradans 2-40	3.2.1.1	23027235	1E-140	13507463	1245	414	643	57

FIGURE 35J

402, 403	ALPHA-AMYLASE PRECURSOR (1,4-ALPHA-D-GLUCAN GLUCANOHYDROLASE)	Pseudoalteromonas haloplanktis	3.2.1.1	6226551	0	2467084	1995	664	669	66
404, 405	hypothetical protein [Microbulbifer degradans 2-40]	Microbulbifer degradans 2-40	3.2.1.1	23027244	1E-155	2337886	1653	550	566	48
406, 407	amylase precursor	Aeromonas hydrophila	3.2.1.1	141870	0	141869	1398	465	464	94
408, 409	alpha-amylase [Xanthomonas axonopodis pv. citri str. 306]	Xanthomonas axonopodis pv. citri str. 306	3.2.1.1	21106921	0	155351	1476	491	475	64
410, 411	hypothetical protein [Microbulbifer degradans 2-40]	Microbulbifer degradans 2-40	3.2.1.1	23027235	0	14861204	1875	624	643	59
412, 413	hypothetical protein [Chloroflexus aurantiacus]	Chloroflexus aurantiacus	3.2.1.1	22971473	0	8250619	2088	695	627	53
414, 415	hypothetical protein [Nostoc punctiforme]	Nostoc punctiforme	3.2.1.1	23126762	3E-38	14091925	2262	753	552	25
416, 417	beta-agarase [Pseudomonas sp. W7]	Pseudomonas sp. W7	3.2.1.1	9081816	5E-31	14518450	1344	447	642	27
418, 419	alpha-amylase [Xanthomonas campestris pv. campestris]	Xanthomonas campestris pv. campestris	3.2.1.1	19224331	1E-169	155351	1455	484	475	60

FIGURE 35K

420, 421	hypothetical protein [Microbulbifer degradans 2-40]	Microbulbifer degradans 2-40	3.2.1.1	23027235	1E-149	18899	1308	435	643	57
422, 423	secreted alpha-amylase, [Streptomyces coelicolor A3(2)]	Streptomyces coelicolor A3(2)	3.2.1.1	21225304	4E-87	6855156	2751	916	993	14
424, 425	cyclomaltodextrin glucanotransferase (EC 2.4.1.19) precursor [validated] - Bacillus circulans (strain 8)	Bacillus circulans	2.4.1.19	279549	0	39565	2169	722	718	80
426, 427	alpha-amylase, Alpha amylase, catalytic domain [Bacillus anthracis A2012]	Bacillus anthracis str. A2012	3.2.1.41	21400626	1E-151	10728478	2139	712	724	42
428, 429	alpha-amylase [Xanthomonas campestris pv. campestris]	Xanthomonas campestris pv. campestris	3.2.1.1	19224331	1E-142	9789644	1425	474	475	52
430, 431	alpha-amylase [Bacillus sp. TS-23]	Bacillus sp. TS-23	3.2.1.1	722279	0	722278	1650	550	613	85
432, 433	alpha-amylase (EC 3.2.1.1) precursor - Streptomyces violaceus	Streptomyces violaceus	3.2.1.1	80864	0	7619766	1752	583	569	64

FIGURE 35L

434, 435	alpha-amylase (EC 3.2.1.1) isozyme III - rice	Oryza sativa	3.2.1.1	11263719	1E-110	3769329	1302	433	437	48

FIGURE 35M

438, 439	hypothetical protein [Chloroflexus aurantiacus]	22970478	1E-166	Chloroflexus aurantiacus	Thermus flavus amyloma ltase encoding DNA.	AAW83330	1E-157	Thermus rubens glucanotransferase gene amplifying ruben-Xba PCR primer.	AAD04867	4E-04	2.4.1.25	1539	512	1806	601	54	58
440, 441	Crystal Structure Of Amyk38 N289h Mutant	34811325	6E-19		Alpha-amylase K38AMY mutagenic PCR primer Q209.	AAO21008	8E-20	Shewanella sp. SCRC-21406 (FERM BP-5979) ORFb DNA SEQ ID NO:4.	AAA92470	0.87	3.2.1.1	3447	1148	7875	480	13	
442, 443	GLUCAN 1,4-ALPHA-MALTOTETRAHYDROLASE PRECURSOR, MALTOTETRAOSE-FORMING AMYLASE, MALTOTETRAOSE-FORMING EXO-AMYLASE	2506188	0	Pseudomonas stutzeri	Maltotetraose.	AAR07282	0	Maltotetraose.	AAQ05095	0	3.2.1.1	1668	555		548	75	78

FIGURE 35N

444, 445	alpha-amylase Thermoactinomyces vulgaris.	322327	1E-107	Thermoactinomyces vulgaris	Alpha amylase DNA PCR primer #1.	ABU03092	1E-121	Human cDNA encoding secreted/transmembrane protein #76.	ABX16619	0.38	2.4.1.19	1521	506	2196	502	44	43
446, 447	alpha-amylase Thermoactinomyces vulgaris.	322327	5E-74	Thermoactinomyces vulgaris	Alpha amylase DNA PCR primer #1.	ABU03092	3E-56	NOVX related reverse PCR primer SEQ ID No 149.	ABT16037	0.89	2.4.1.19	3537	1178		482	14	35
448, 449	AmyM [uncultured bacterium].	37183425	1E-158	uncultured bacterium	Alpha amylase DNA PCR primer #1.	ABU03133	1E-126	Human cDNA SEQ ID NO 201.	ABQ93381	0.099	2.4.1.25	1575	524	1554	517	51	57
450, 451	alpha-amylase 3 [Bacteroides thetaiotaomicron VPI-5482]	29346183	1E-151	Bacteroides thetaiotaomicron VPI-5482	Alpha amylase DNA PCR primer #1.	ABU03134	4E-16	cDNA encoding human transporter polypeptide.	AAS16905	0.47		1890	629		565	43	49
452, 453	AmyA [uncultured bacterium]	37222142	1E-135	uncultured bacterium	Alpha amylase DNA PCR primer #1.	ABU03131	1E-72	Alpha amylase DNA PCR primer #1.	ABX08502	0.007	3.2.1.1	1734	577		608	43	56

FIGURE 35O

454, 455	alpha-amylase [Hordeum vulgare].	295804	1E-116	Hordeum vulgare	Endoxyloglucan transferase sequence #165.	AAM00108	1E-118	Partial sequence of tomato Ca2+ ATPase.	AAZ29771	0.33	3.2.1.1	1341	446	6263	430	47	53
456, 457	AmyA [uncultured bacterium]	37222142	1E-136	uncultured bacterium	Alpha amylase DNA PCR primer #1.	ABU03131	2E-73	Alpha amylase DNA PCR primer #1.	ABX08502	0.028	3.2.1.1	1734	577		608	44	56
458, 459	ALPHA-AMYLASE PRECURSOR (1,4-ALPHA-D-GLUCAN GLUCANOHYDROLASE).	113814	3E-81	Bacillus megaterium	Alpha amylase DNA PCR primer #1.	ABU03115	7E-85	Alpha amylase DNA PCR primer #1.	ABX08486	0.002	3.2.1.1	1698	565	1620	539	30	47
460, 461	maltase [Aspergillus oryzae].	14278921	1E-133	Aspergillus oryzae	Vibrio harveyi endoglucanase DNA.	AAW34990	1E-108	Chromosome 13q31-q33 biallelic marker containing amplicon SEQ ID 182.	AAH51601	0.11	3.2.1.20	1752	583	1725	574	43	53

FIGURE 35P

462-466	related to glucoamylase precursor [Neurospora crassa]	38524238	1E-143	Neurospora crassa	Thielavia terrestris glucoamylase DNA PCR primer #4.	AAM51596	2E-39	Aspergillus oryzae polynucleotide SEQ ID NO 3150.	ABZ54259	0.001		1206	401		405	60	63
467-474	hypothetical protein MG03287.4 [Magnaporthe grisea 70-15]	38105244	1E-177	Magnaporthe grisea 70-15	Sequence of amylase gene and upstream regulatory DNA.	AAR09359	1E-125	Alpha amylase DNA PCR primer #1.	ABX08477	0.61	3.2.1.1	2421	806		518	36	39
475-479	putative alpha-1,3-glucan synthase [Aspergillus fumigatus].	16418019	0	Aspergillus fumigatus	Maize Starch synthase IIb (SSIIb) LINKR domain related protein #42.	ABU06520	3E-13	Aspergillus oryzae polynucleotide SEQ ID NO 3150.	ABZ53554	5E-04	2.4.1.183	7131	2376	7281	2426	59	59
480-485	hypothetical protein MG10209.4 [Magnaporthe grisea 70-15]	38101134	1E-175	Magnaporthe grisea 70-15	Amino acid sequence of a fungamyl-like alpha-amylase.	AAB84206	1E-103	Fusarium venenatum EST SEQ ID NO:1176.	AAF07664	5E-65	3.2.1.1	1383	460		600	62	62

FIGURE 35Q

486-493	GLUCOAMYLASE P PRECURSOR (GLUCAN 1,4-ALPHA-GLUCOSIDASE) (1,4-ALPHA-D GLUCAN GLUCOHYDROLASE).	461509	0	Amorphotheca resinae	cDNA encoding glucoamylase P.	AAW30155	0	cDNA encoding glucoamylase P.	AAT90830	5E-04	3.2.1.3	1932	643		616	61	60
494-499	alpha-amylase AmyA [Emericella nidulans].	6561867	1E-144	Emericella nidulans	Mutant alpha-amylase.	AAR46065	1E-142	Fusarium venenatum EST SEQ ID NO:1176.	AAF12832	4E-04	3.2.1.1	1479	492	1473	490	49	54
500-510	alpha-glucosidase [Schizosaccharomyces pombe].	19111855	1E-145	Schizosaccharomyces pombe	PCR primer B1 from J092340 81.	AAW27300	1E-123	Drosophila melanogaster polypeptide SEQ ID NO 24465.	ABL18081	0.002	3.2.1.20	1833	610	1740	579	42	50
511-516	acid-stable alpha-amylase [Aspergillus kawachii].	2570150	1E-125	Aspergillus kawachii	Amino acid sequence of a fungamyl like alpha-amylase.	AAB84206	1E-122	Arabidopsis thaliana stress regulated gene SEQ ID NO 1888.	ABZ16178	1.6	3.2.1.1	1650	549	1923	640	43	51

FIGURE 35R

517, 518	alpha-amylase AmyA [Emericella nidulans].	6561867	1E-110	Emericella nidulans	Alpha-amylase variant with leucine at position 84.	AAR24136	1E-110	S. pneumoniae type 4 strain protein from coding region #1864.	ABS56454	0.44	3.2.1.1	1758	585	1473	490	37	42
519-523	alpha-glucosidase [Aspergillus oryzae].	23503475	1E-179	Aspergillus oryzae	PCR primer B1 from J092340 81.	AAW27300	1E-129	Fusarium venenatum EST SEQ ID NO:1176.	AAF08465	0.002	3.2.1.20	1788	595	1809	602	50	56
524-528	OLIGO-1,6-GLUCOSIDASE (SUCRASE-ISOMALTASE) (LIMIT DEXTRINASE) (ISOMALTASE) (DEXTRIN 6-ALPHA-D-GLUCANOHYDROLASE).	129007	1E-145	Geobacillus thermoglucosidasius	PCR primer B1 from J092340 81.	AAW27300	1E-138	Bacillus licheniformis genomic sequence tag (GST) #933.	ABK73353	0.007	3.2.1.10	1755	584		562	46	51
529-533	alpha-glucosidase [Aspergillus oryzae].	23503475	1E-165	Aspergillus oryzae	PCR primer B1 from J092340 81.	AAW27300	1E-141	Aspergillus oryzae polynucleotide SEQ ID NO 3150.	ABZ54241	5E-07	3.2.1.20	1797	598	1809	602	48	55

FIGURE 35S

534-540	hypothetical protein MG10209.4 [Magnaporthe grisea 70-15]	38101134	0	Magnaporthe grisea 70-15	Amino acid sequence of a fungamyl-like alpha-amylase.	AAB84206	1E-105	Fusarium venenatum EST SEQ ID NO:1176.	AAF07664	5E-13	3.2.1.1	1770	589		600	65	68
541-545	hypothetical protein [Neurospora crassa] gb|EAA32559.1	32411795	0	Neurospora crassa	HIV multifunctional fusion polypeptide.	AAR13230	1E-137	HIV multifunctional fusion polypeptide.	AAQ12770	1E-04	3.2.1.20	1992	663		608	60	
546-553	glucan 1,4-alpha-glucosidase (EC 3.2.1.3) precursor - Neurospora crassa.	486943	0	Neurospora crassa	Thielavia terrestris glucoamylase DNA PCR primer #4.	AAM51596	0	Thielavia terrestris glucoamylase DNA PCR primer #4.	ABA01139	5E-07	3.2.1.3	2019	672		626	58	36
554-559	maltase [Aspergillus oryzae].	14278921	0	Aspergillus oryzae	HIV multifunctional fusion polypeptide.	AAR13230	1E-147	Fusarium venenatum EST SEQ ID NO:1176.	AAF13291	5E-04	3.2.1.20	1797	598	1725	574	57	59

FIGURE 35T

560-566	ALPHA-AMYLASE A PRECURSOR (1,4-ALPHA-D-GLUCAN GLUCANOHYDROLASE A).	1703298	1E-125	Aspergillus niger var. awamori	Amino acid sequence of a fungamyl like alpha-amylase.	AAB84206	1E-125	Vector pPR70-4 xlnB expression element region.	AAV61459	0.38	3.2.1.1	1524	507		498	46	
567, 568	beta-agarase [Pseudomonas sp. W7].	9081816	1E-126	Pseudomonas sp. W7	Alpha amylase. DNA PCR primer #1.	ABU03140	1E-113	Human secreted protein SEQ ID NO 792.	ABZ66894	1.4	3.2.1.1	1446	481	1926	642	49	57
569, 570		40189607	0									1095	364	1095	364	100	100
571, 572	alpha-amylase [Bacillus sp. TS-23].	722279	0	Bacillus sp. TS-23	Alpha amylase DNA PCR primer #1.	ABU03102	0	Alpha amylase DNA PCR primer #1.	ABX08491	3E-11	3.2.1.1	1839	612	1842	613	59	56
573, 574	CYCLOMALTODEXTRIN GLUCANOTRANSFERASE PRECURSOR (CYCLODEXTRIN-GLYCOSYLTRANSFERASE) (CGTASE).	1351937	0	Thermoanaerobacterium thermosulfurigenes	Thermoanaerobacter CGTase variant G180S.	AAW06772	0	Thermoanaerobacter CGTase variant G180S.	AAT44121	0	2.4.1.19	2133	710		710	100	79

FIGURE 35U

575, 576	alpha-amylase [Bacillus sp. TS-23].	722279	7E-77	Bacillus sp. TS-23	Nucleotide sequence of an alpha-amylase of Bacillus NCIMB 40916.	AAB67655	0	Nucleotide sequence of an alpha-amylase of Bacillus NCIMB 40916.	AAF55662	0	3.2.1.1	1677	558	1764	587	99	97
577, 578	cyclodextrin glucanosyltransferase [Bacillus circulans].	11139208	0	Bacillus circulans	A cyclomaltodextrin glucanotransferase (CTGase) variant.	AAY02599	0	Plasmid pTN603 encoding novel amylase gene.	AAN60705	2E-10	2.4.1.19	2160	719	2142	713	67	66
579, 580	AMYLOPULLULANASE PRECURSOR (ALPHA-AMYLASE/PULLULANASE)	114076	0	Thermoanaerobacter thermohydrosulfuricus	Alpha amylase pullulanase gene.	AAR08221	0	Alpha amylase pullulanase gene.	AAQ06827	0	3.2.1.135	5010	1669		1475	87	90

FIGURE 35V

581, 582	CYCLOMALTODEXTRIN GLUCANOTRANSFERASE PRECURSOR (CYCLODEXTRIN-GLYCOSYLTRANSFERASE) (CGTASE).	399222	0	Bacillus sp. 17-1	Cyclomaltodextrin glucotransferase 17-1 gene.	AAR10052	0	Sequence encoding cyclomaltodextrin glucanotransferase (CGTase) having alkaline optimum pH.	AAQ01810	0	2.4.1.19	2142	713		713	95	89
583, 584	amylopullulanase (Geobacillus stearothermophilus).	12006232	0	Geobacillus stearothermophilus	Alpha amylase pullulanase gene.	AAR08221	0	Mycobacterium tuberculosis strain H37Rv genome SEQ ID NO 2.	AAI99683	2E-06	3.2.1.20	6618	2205	6057	2018	29	46
585, 586	alpha-amylase [Bacillus sp. TS-23].	722279	0	Bacillus sp. TS-23	Alpha amylase DNA PCR primer #1.	ABU03091	0	Alpha amylase DNA PCR primer #1.	ABX08462	0	3.2.1.1	1860	619	1842	613	82	75
603, 604	alpha-amylase [Bacillus sp. TS-23].	722279	1E-77	Bacillus sp. TS-23	Nucleotide sequence of an alpha-amylase of Bacillus NCIMB 40916.	AAB67655	0	Nucleotide sequence of an alpha-amylase of Bacillus NCIMB 40916.	AAF55662	0	3.2.1.1	1773	590	1764	587	97	95

FIGURE 35W

605, 606	CYCLOMALTODEXTRIN GLUCANOTRANSFERASE PRECURSOR (CYCLODEXTRIN-GLYCOSYLTRANSFERASE) (CGTASE).	399219	0	Bacillus circulans	A cyclomaltodextrin glucanotransferase (CTGase) variant.	AAY02599	0	Plasmid pTN603 encoding novel amylase gene.	AAN60705	8E-65	2.4.1.19	2160	719		718	74	70
607, 608	amylopullulanase [Geobacillus stearothermophilus].	12006232	0	Geobacillus stearothermophilus	Alpha amylase DNA PCR primer #1.	ABU03083	0	Alpha amylase DNA PCR primer #1.	ABX08454	0	3.2.1.135	3708	1235	6057	2018	77	71
609, 610	alpha-amylase [Bacillus sp. TS-23].	722279	7E-77	Bacillus sp. TS-23	Nucletoide sequence of an alpha-amylase of Bacillus NCIMB 40916.	AAB67655	0	Nucletoide sequence of an alpha-amylase of Bacillus NCIMB 40916.	AAF55662	0	3.2.1.1	1764	587	1764	587	98	97
611, 612	hypothetical protein - Bacillus sp.	2126830	0	Bacillus sp.	Alpha amylase DNA PCR primer #1.	ABU03083	0	Alpha amylase DNA PCR primer #1.	ABX08454	0	3.2.1.135	3807	1268		2032	75	

FIGURE 35X

613, 614	alpha-amylase [Bacillus sp. TS-23].	722279	3E-77	Bacillus sp. TS-23	Nucletoide sequence of an alpha-amylase of Bacillus NCIMB 40916.	AAB67655	0	Nucletoide sequence of an alpha-amylase of Bacillus NCIMB 40916.	AAF55662	0	3.2.1.1	1764	587	1764	587	99	97
615, 616	ALPHA-AMYLASE PRECURSOR (1,4-ALPHA-D-GLUCAN GLUCANOHYDROLASE).	113822	0	Streptomyces violaceus	Amylase from Streptomyces griseus IMRU 3570.	AAR08263	0	Amylase from Streptomyces limolus.	AAN80309	2E-61	3.2.1.1	1752	583		569	64	72
617, 618	GLUCAN 1,4-ALPHA-MALTOTETRAHYDROLASE PRECURSOR (G4-AMYLASE) (MALTOTETRAOSE-FORMING AMYLASE) (EXO-MALTOTETRAOHYDROLASE) (MALTOTETRAOSE-FORMING EXO-AMYLASE).	113760	0	Pseudomonas saccharophila	Maltotetraose.	AAR07282	0	Maltotetraose.	AAQ06095	0	3.2.1.1	1668	555		551	74	78

FIGURE 35Y

619, 620	ALPHA-AMYLASE PRECURSOR (1,4-ALPHA-D-GLUCAN GLUCANOHYDROLASE).	113822	0	Streptomyces violaceus	Amylase from Streptomyces griseus IMRU 3570.	AAR08263	0	Amylase from Streptomyces griseus IMRU 3570.	AAQ06844	0	3.2.1.1	1716	571		569	91	81
621, 622	alpha-amylase [Bacillus sp. TS-23].	722279	3E-77	Bacillus sp. TS-23	Nucletolde sequence of an alpha-amylase of Bacillus NCIMB 40916.	AAB67655	0	Nucletoide sequence of an alpha-amylase of Bacillus NCIMB 40916.	AAF55662	0	3.2.1.1	1773	590	1764	587	98	97

SEQUENCE LISTING

<110> Diversa Corporation
Callen, Walter
Richardson, Toby
Frey, Gerhard
Gray, Kevin A.
Kerovuo, Janne S.
Slupska, Malgorzata
Barton, Nelson
O'Donoghue, Eileen

<120> Amylases, Nucleic Acids Encoding Them and Methods for Making and Using Them

<130> 564462006151

<150> 60/459,014
<151> 2003-03-28

<150> 10/385,305
<151> 2003-03-06

<160> 622

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 1311
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetically generated oligonucleotide

<400> 1
atggcaaagt attccgagct cgaagagggc gggctcataa tgcaggcctt ctactgggac 60
gtccccatgg gaggaatctg gtgggacacg atagcccaga agatacccga ctgggcaagc 120
gccgggattt cggcgatatg gattcccccg gcgagcaagg gcatgggcgg cgcctattcg 180
atgggctacg acccctacga cttctttgac ctcggtgagt acgaccagaa gggaacggta 240
gagacgcgct ttggctccaa gcaggagctc gtgaacatga taaacaccgc ccacgcctat 300
ggcatgaagg taatagccga tatagtcatc aaccaccgcg ccggcggtga cctggagtgg 360
aacccttcg tgaacgacta tacctggacc gacttctcaa aggtcgcgtc gggtaaatac 420
acggccaact acctcgactt ccacccgaac gagctccatg cgggcgattc cggaacattt 480
ggaggctatc ccgacatatg ccacgacaag agctgggacc agtactggct ctgggccagc 540
caggagagct acgcggcata tctcaggagc atcggcatcg atgcctggcg cttcgactac 600
gtcaagggct acggagcgtg ggtcgtcaag gactggctgg actggtgggg aggctgggcc 660
gtcggggagt actgggacac aaacgttgat gcactgctca actgggccta ctcgagcgat 720
gcaaaagtct tcgacttccc gctctactac aagatggacg cggcctttga caacaagaac 780
attcccgcac tcgtcgaggc cctcaagaac gggggcacag tcgtcagccg cgacccgttt 840
aaggccgtaa ccttcgttgc aaaccacgac accgatataa tctggaacaa gtatccagcc 900
tacgcgttca tcctcaccta cgagggccag ccgacaatat tctaccgcga ctacgaggag 960
tggctcaaca aggataagct caagaacctc atctggatac atgacaacct cgccggagga 1020
agcactgaca tcgtttacta cgacaacgac gagctgatat tcgtgagaaa cggctacgga 1080
agcaagccgg gactgataac atacatcaac ctcgcctcaa gcaaagccgg aaggtgggtt 1140
tacgttccga agttcgcagg ctcgtgcata cacgagtaca ccggcaatct cggcggctgg 1200
gtggacaagt gggtggactc aagcggctgg gtctacctcg aggctcctgc ccacgacccg 1260
gccaacggcc agtacggcta ctccgtctgg agctactgcg gtgttgggtg a 1311

<210> 2
<211> 436
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetically generated peptide

<400> 2
Met Ala Lys Tyr Ser Glu Leu Glu Glu Gly Gly Leu Ile Met Gln Ala
 1   5   10   15
Phe Tyr Trp Asp Val Pro Met Gly Gly Ile Trp Trp Asp Thr Ile Ala
            20   25   30
Gln Lys Ile Pro Asp Trp Ala Ser Ala Gly Ile Ser Ala Ile Trp Ile
        35   40   45
Pro Pro Ala Ser Lys Gly Met Gly Gly Ala Tyr Ser Met Gly Tyr Asp
    50   55   60
Pro Tyr Asp Phe Phe Asp Leu Gly Glu Tyr Asp Gln Lys Gly Thr Val
65   70   75   80
Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Val Asn Met Ile Asn Thr
                85   90   95
Ala His Ala Tyr Gly Met Lys Val Ile Ala Asp Ile Val Ile Asn His
            100   105   110
Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Asn Asp Tyr Thr
        115   120   125
Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr
    130   135   140
Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe
145   150   155   160
Gly Gly Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp
                165   170   175
Leu Trp Ala Ser Gln Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly
            180   185   190
Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp Val
        195   200   205
Val Lys Asp Trp Leu Asp Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr
    210   215   220
Trp Asp Thr Asn Val Asp Ala Leu Leu Asn Trp Ala Tyr Ser Ser Asp
225   230   235   240
Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Ala Ala Phe
                245   250   255
Asp Asn Lys Asn Ile Pro Ala Leu Val Glu Ala Leu Lys Asn Gly Gly
            260   265   270
Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn
        275   280   285
His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile
    290   295   300
Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu
305   310   315   320
Trp Leu Asn Lys Asp Lys Leu Lys Asn Leu Ile Trp Ile His Asp Asn
                325   330   335
Leu Ala Gly Gly Ser Thr Asp Ile Val Tyr Tyr Asp Asn Asp Glu Leu
            340   345   350
Ile Phe Val Arg Asn Gly Tyr Gly Ser Lys Pro Gly Leu Ile Thr Tyr
        355   360   365
Ile Asn Leu Ala Ser Ser Lys Ala Gly Arg Trp Val Tyr Val Pro Lys
    370   375   380
Phe Ala Gly Ser Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp
385   390   395   400
Val Asp Lys Trp Val Asp Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro
                405   410   415
Ala His Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr
            420   425   430
Cys Gly Val Gly
        435

<210> 3
<211> 1311
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetically generated oligonucleotide

<400> 3
atggccaagt acctggagct cgaagagggc gggctcataa tgcaggcctt ctactgggac 60

```
gtccccatgg gaggaatctg gtgggacacg atagcccaga agatacccga ctgggcaagc     120
gccgggattt cggcgatatg gattcccccg gcgagcaagg gcatgggcgg cgcctattcg     180
atgggctacg acccctacga cttctttgac ctcggtgagt acgaccagaa gggaacggta     240
gagacgcgct ttggctccaa gcaggagctc gtgaacatga taaacaccgc ccacgcctac     300
ggcatcaagg tcatcgcaga catagtaatc aaccaccgcg ccggaggaga ccttgagtgg     360
aacccttcg tcaatgacta cacctggacg gacttctcga aggtcgcttc cggcaagtac     420
acggccaatt acctcgactt ccacccgaac gagctccatg cgggcgattc cggaacattt     480
ggaggctatc ccgacatatg ccacgacaag agctgggacc agtactggct ctgggccagc     540
caggagagct acgcggcata tctcaggagc atcggcatcg atgcctggcg cttcgactac     600
gtcaagggct atgctccctg ggtcgtcaag gactggctga actggtgggg aggctgggcg     660
gttggagagt actgggacac caacgtcgac gctgttctca actgggcata ctcgagcggt     720
gccaaggtct ttgacttcgc cctctactac aagatggatg aggcctttga caacaaaaac     780
attccagcgc tcgtctctgc ccttcagaac ggccagactg ttgtctcccg cgacccgttc     840
aaggccgtaa cctttgtagc aaaccacgac accgatataa tctggaacaa gtatccagcc     900
tacgcgttca tcctcaccta cgagggccag ccgacaatat tctaccgcga ctacgaggag     960
tggctcaaca aggataagct caagaacctc atctggatac atgacaacct cgccggagga    1020
agcactgaca tcgtttacta cgacaacgac gagctgatat tcgtgagaaa cggctacgga    1080
agcaagccgg gactgataac atacatcaac ctcgcctcaa gcgaagccgg aaggtgggtc    1140
tacgttccga agttcgcggg agcgtgcatc cacgagtaca ccggcaacct cggcggctgg    1200
gtggacaagt gggtggactc aagcgggtgg gtgtacctcg aggcccctgc ccacgacccg    1260
gccaacggct attacggcta ctccgtctgg agctattgcg gtgttgggtg a             1311

<210> 4
<211> 436
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetically generated peptide

<400> 4
Met Ala Lys Tyr Leu Glu Leu Glu Glu Gly Gly Leu Ile Met Gln Ala
 1               5                  10                  15
Phe Tyr Trp Asp Val Pro Met Gly Gly Ile Trp Trp Asp Thr Ile Ala
            20                  25                  30
Gln Lys Ile Pro Asp Trp Ala Ser Ala Gly Ile Ser Ala Ile Trp Ile
        35                  40                  45
Pro Pro Ala Ser Lys Gly Met Gly Gly Ala Tyr Ser Met Gly Tyr Asp
    50                  55                  60
Pro Tyr Asp Phe Phe Asp Leu Gly Glu Tyr Asp Gln Lys Gly Thr Val
65                  70                  75                  80
Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Val Asn Met Ile Asn Thr
                85                  90                  95
Ala His Ala Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His
            100                 105                 110
Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Asn Asp Tyr Thr
        115                 120                 125
Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr
    130                 135                 140
Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe
145                 150                 155                 160
Gly Gly Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp
                165                 170                 175
Leu Trp Ala Ser Gln Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly
            180                 185                 190
Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Ala Pro Trp Val
        195                 200                 205
Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr
    210                 215                 220
Trp Asp Thr Asn Val Asp Ala Val Leu Asn Trp Ala Tyr Ser Ser Gly
225                 230                 235                 240
Ala Lys Val Phe Asp Phe Ala Leu Tyr Tyr Lys Met Asp Glu Ala Phe
                245                 250                 255
Asp Asn Lys Asn Ile Pro Ala Leu Val Ser Ala Leu Gln Asn Gly Gln
            260                 265                 270
Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn
        275                 280                 285
His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile
```

290 295 300
Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu
305 310 315 320
Trp Leu Asn Lys Asp Lys Leu Lys Asn Leu Ile Trp Ile His Asp Asn
325 330 335
Leu Ala Gly Gly Ser Thr Asp Ile Val Tyr Tyr Asp Asn Asp Glu Leu
340 345 350
Ile Phe Val Arg Asn Gly Tyr Gly Ser Lys Pro Gly Leu Ile Thr Tyr
355 360 365
Ile Asn Leu Ala Ser Ser Glu Ala Gly Arg Trp Val Tyr Val Pro Lys
370 375 380
Phe Ala Gly Ala Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp
385 390 395 400
Val Asp Lys Trp Val Asp Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro
405 410 415
Ala His Asp Pro Ala Asn Gly Tyr Tyr Gly Tyr Ser Val Trp Ser Tyr
420 425 430
Cys Gly Val Gly
435

<210> 5
<211> 1311
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetically generated oligonucleotide

<400> 5
atggccaagt actccgagct ggaagagggc ggcgttataa tgcaggcctt ctactgggac 60
gtcccaggtg gaggaatctg gtgggacacc atcaggagca agataccgga gtggtacgag 120
gcgggaatat ccgccatttg gattcccccg gcaagcaagg gcatgggcgg cgcctattcg 180
atgggctacg acccctacga cttctttgac ctcggtgagt acgaccagaa gggaacggta 240
gagacgcgct ttggctccaa gcaggagctc gtgaacatga taaacaccgc ccacgcctat 300
ggcatgaagg taatagccga tatagtcatc aaccaccgcg ccggcggtga cctggagtgg 360
aacccttcg tgaacgacta tacctggacc gacttctcaa aggtcgcgtc gggtaaatac 420
acggccaact acctcgactt ccacccgaac gagctccatg cgggcgattc cggaacattt 480
ggaggctatc ccgacatatg ccacgacaag agctgggacc agtactggct ctgggccagc 540
caggagagct acgcggcata tctcaggagc atcggcatcg atgcctggcg cttcgactac 600
gtcaagggct atgctccctg ggtcgtcaag gactggctga actggtgggg aggctgggcg 660
gttggagagt actgggacac caacgtcgac gctgttctca actgggcata ctcgagcggt 720
gccaaggtct ttgacttcgc cctctactac aagatggatg aggcctttga caacaaaaac 780
attccagcgc tcgtctctgc ccttcagaac ggccagactg ttgtctcccg cgacccgttc 840
aaggccgtaa cctttgtagc aaaccacgac accgatataa tctggaacaa gtaccttgct 900
tatgctttca tcctcaccta cgaaggccag cccgtcatat tctaccgcga ccacgaggag 960
tggctcaaca aggacaggtt gaacaacctc atatggatac acgaccacct cgcaggtgga 1020
agcaccgaca tagtctacta cgataacgat gaactcatct tcgtcaggaa cggctacggg 1080
gacaagccgg ggcttataac ctacatcaac ctaggctcga gcaaggccgg aaggtgggtt 1140
tatgtgccga agttcgcggg cgcgtgcatc cacgagtata ctggtaacct cggaggctgg 1200
gtagacaagt acgtctactc aagcggctgg gtctatctcg aagctccagc ttacgaccct 1260
gccaacgggc agtatggcta ctccgtgtgg agctactgcg gggtgggctg a 1311

<210> 6
<211> 436
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetically generated peptide

<400> 6
Met Ala Lys Tyr Ser Glu Leu Glu Glu Gly Gly Val Ile Met Gln Ala
1 5 10 15
Phe Tyr Trp Asp Val Pro Gly Gly Gly Ile Trp Trp Asp Thr Ile Arg
20 25 30
Ser Lys Ile Pro Glu Trp Tyr Glu Ala Gly Ile Ser Ala Ile Trp Ile
35 40 45
Pro Pro Ala Ser Lys Gly Met Gly Gly Ala Tyr Ser Met Gly Tyr Asp

```
    50                  55                  60
Pro Tyr Asp Phe Phe Asp Leu Gly Glu Tyr Asp Gln Lys Gly Thr Val
65                  70                  75                  80
Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Val Asn Met Ile Asn Thr
                85                  90                  95
Ala His Ala Tyr Gly Met Lys Val Ile Ala Asp Ile Val Ile Asn His
            100                 105                 110
Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Asn Asp Tyr Thr
        115                 120                 125
Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr
    130                 135                 140
Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe
145                 150                 155                 160
Gly Gly Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp
                165                 170                 175
Leu Trp Ala Ser Gln Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly
            180                 185                 190
Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Ala Pro Trp Val
        195                 200                 205
Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr
    210                 215                 220
Trp Asp Thr Asn Val Asp Ala Val Leu Asn Trp Ala Tyr Ser Ser Gly
225                 230                 235                 240
Ala Lys Val Phe Asp Phe Ala Leu Tyr Tyr Lys Met Asp Glu Ala Phe
                245                 250                 255
Asp Asn Lys Asn Ile Pro Ala Leu Val Ser Ala Leu Gln Asn Gly Gln
            260                 265                 270
Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn
        275                 280                 285
His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Leu Ala Tyr Ala Phe Ile
    290                 295                 300
Leu Thr Tyr Glu Gly Gln Pro Val Ile Phe Tyr Arg Asp His Glu Glu
305                 310                 315                 320
Trp Leu Asn Lys Asp Arg Leu Asn Asn Leu Ile Trp Ile His Asp His
                325                 330                 335
Leu Ala Gly Gly Ser Thr Asp Ile Val Tyr Tyr Asp Asn Asp Glu Leu
            340                 345                 350
Ile Phe Val Arg Asn Gly Tyr Gly Asp Lys Pro Gly Leu Ile Thr Tyr
        355                 360                 365
Ile Asn Leu Gly Ser Ser Lys Ala Gly Arg Trp Val Tyr Val Pro Lys
    370                 375                 380
Phe Ala Gly Ala Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp
385                 390                 395                 400
Val Asp Lys Tyr Val Tyr Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro
                405                 410                 415
Ala Tyr Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr
            420                 425                 430
Cys Gly Val Gly
        435

<210> 7
<211> 16
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 7
Met Ala Leu Glu Glu Gly Gly Leu Ile Met Gln Ala Phe Tyr Trp Asp
 1               5                  10                  15

<210> 8
<211> 26
<212> PRT
<213> Bacterial

<400> 8
```

Met Lys Pro Ala Lys Leu Leu Val Phe Val Leu Val Val Ser Ile Leu
 1   5   10   15
Ala Gly Leu Tyr Ala Gln Pro Ala Gly Ala
            20   25

<210> 9
<211> 1311
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetically generated oligonucleotide

<400> 9
atggccaagt actccgagct ggaagagggc gggctcataa tgcaggcctt ctactgggac 60
gtccccatgg gaggaatctg gtgggacacg atagcccaga agatacccga ctgggcaagc 120
gccgggattt cggcgatatg gattcccccg gcgagcaagg gcatgggcgg cgcctattcg 180
atgggctacg acccctacga cttctttgac ctcggtgagt acgaccagaa gggaacggta 240
gagacgcgct ttggctccaa gcaggagctc gtgaacatga taaacacggc ccatgcctac 300
ggcataaagg tcatagcgga catcgtcata aaccaccgcg caggcggaga cctcgagtgg 360
aacccgttcg ttggggacta cacctggacg gacttctcaa aggtggcctc gggcaaatat 420
actgccaact acctcgactt ccacccgaac gagctccatg cgggcgattc cggaacattt 480
ggaggctatc ccgacatatg ccacgacaag agctgggacc agtactggct ctgggccagc 540
caggagagct acgcggcata tctcaggagc atcggcatcg atgcctggcg cttcgactac 600
gtcaagggct atgctccctg ggtcgtcaag gactggctga actggtgggg aggctgggcg 660
gttggagagt actgggacac caacgtcgac gctgttctca actgggcata ctcgagcggt 720
gccaaggtct ttgacttcgc cctctactac aagatggacg aggccttcga taacaacaac 780
attcccgccc tggtggacgc cctcagatac ggtcagacag tggtcagccg cgacccgttc 840
aaggctgtga cgtttgtagc caaccacgat accgatataa tctggaacaa gtatccagcc 900
tacgcgttca tcctcaccta cgagggccag ccgacaatat tctaccgcga ctacgaggag 960
tggctcaaca aggataagct caagaacctc atctggatac atgacaacct cgccggagga 1020
agcactgaca tcgtttacta cgacaacgac gagctgatat tcgcgagaaa cggctacgga 1080
agcaagccgg gactgataac atacatcaac ctcgcctcaa gcaaagccgg aaggtgggtt 1140
tacgttccga agttcgcagg ctcgtgcata cacgagtaca ccggcaatct cggcggctgg 1200
gtggacaagt gggtggactc aagcggctgg gtctacctcg aggctcctgc ccacgacccg 1260
gccaacggcc agtacggcta ctccgtctgg agctactgcg gtgttgggtg a 1311

<210> 10
<211> 436
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetically generated peptide

<400> 10
Met Ala Lys Tyr Ser Glu Leu Glu Glu Gly Gly Leu Ile Met Gln Ala
 1   5   10   15
Phe Tyr Trp Asp Val Pro Met Gly Gly Ile Trp Trp Asp Thr Ile Ala
            20   25   30
Gln Lys Ile Pro Asp Trp Ala Ser Ala Gly Ile Ser Ala Ile Trp Ile
        35   40   45
Pro Pro Ala Ser Lys Gly Met Gly Gly Ala Tyr Ser Met Gly Tyr Asp
    50   55   60
Pro Tyr Asp Phe Phe Asp Leu Gly Glu Tyr Asp Gln Lys Gly Thr Val
65   70   75   80
Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Val Asn Met Ile Asn Thr
                85   90   95
Ala His Ala Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His
            100   105   110
Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Gly Asp Tyr Thr
        115   120   125
Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr
    130   135   140
Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe
145   150   155   160
Gly Gly Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp
                165   170   175

```
Leu Trp Ala Ser Gln Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly
            180                 185                 190
Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Ala Pro Trp Val
        195                 200                 205
Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr
    210                 215                 220
Trp Asp Thr Asn Val Asp Ala Val Leu Asn Trp Ala Tyr Ser Ser Gly
225                 230                 235                 240
Ala Lys Val Phe Asp Phe Ala Leu Tyr Tyr Lys Met Asp Glu Ala Phe
                245                 250                 255
Asp Asn Asn Asn Ile Pro Ala Leu Val Asp Ala Leu Arg Tyr Gly Gln
            260                 265                 270
Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn
        275                 280                 285
His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile
    290                 295                 300
Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu
305                 310                 315                 320
Trp Leu Asn Lys Asp Lys Leu Lys Asn Leu Ile Trp Ile His Asp Asn
                325                 330                 335
Leu Ala Gly Gly Ser Thr Asp Ile Val Tyr Tyr Asp Asn Asp Glu Leu
            340                 345                 350
Ile Phe Ala Arg Asn Gly Tyr Gly Ser Lys Pro Gly Leu Ile Thr Tyr
        355                 360                 365
Ile Asn Leu Ala Ser Ser Lys Ala Gly Arg Trp Val Tyr Val Pro Lys
    370                 375                 380
Phe Ala Gly Ser Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp
385                 390                 395                 400
Val Asp Lys Trp Val Asp Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro
                405                 410                 415
Ala His Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr
            420                 425                 430
Cys Gly Val Gly
        435
```

<210> 11
<211> 1311
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetically generated oligonucleotide

<400> 11
```
atggccaagt acctggagct cgaggagggc gggctcataa tgcaggcctt ctactgggac      60
gtccccatgg gaggaatctg gtgggacacg atagcccaga agatacccga ctgggcaagc     120
gccgggattt cggcgatatg gattcccccg gcgagcaagg gcatgggcgg cgcctattcg     180
atgggctacg acccctacga cttctttgac ctcggtgagt acgaccagaa gggaacggta     240
gagacgcgct ttggctccaa gcaggagctc gtgaacatga taaacaccgc ccacgcctat     300
ggcatgaagg taatagccga tatagtcatc aaccaccgcg ccggcggtga cctggagtgg     360
aacccttcg tgaacgacta tacctggacc gacttctcaa aggtcgcgtc gggtaaatac     420
acggccaact acctcgactt ccacccgaac gagctccatg cgggcgattc cggaacattt     480
ggaggctatc ccgacatatg ccacgacaag agctgggacc agtactggct ctgggccagc     540
caggagagct acgcggcata tctcaggagc atcggcatcg atgcctggcg cttcgactac     600
gtcaagggct atgctccctg ggtcgtcaag gactggctga actggtgggg aggctgggcg     660
gttggagagt actgggacac caacgtcgac gctgttctca actgggcata ctcgagcggt     720
gccaaggtct ttgacttcgc cctctactac aagatggacg aggccttcga taacaacaac     780
attcccgccc tggtggacgc cctcagatac ggtcagacag tggtcagccg cgacccgttc     840
aaggctgtga cgtttgtagc caaccacgat accgatataa tctggaacaa gtatccagcc     900
tacgcgttca tcctcaccta cgagggccag ccgacaatat tctaccgcga ctacgaggag     960
tggctcaaca aggatacgct caagaacctc atctggatac atgacaacct cgccggagga    1020
agcacgagca tagtttacta cgacagcgac gagatgatct tcgtgaggaa cggctatgga    1080
agcaagcctg gccttataac ttacatcaac ctcggctcga gcaaggttgg aaggtgggtc    1140
tacgttccga agttcgcggg agcgtgcatc cacgagtaca ccggcaacct cggcggctgg    1200
gtggacaagt gggtggactc aagcgggtgg gtgtacctcg aggcccctgc ccacgacccg    1260
gccaacggct attacggcta ctccgtctgg agctactgcg gtgttggctg a             1311
```

<210> 12

<211> 436
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetically generated peptide

<400> 12
Met Ala Lys Tyr Leu Glu Leu Glu Glu Gly Gly Leu Ile Met Gln Ala
 1   5   10   15
Phe Tyr Trp Asp Val Pro Met Gly Gly Ile Trp Trp Asp Thr Ile Ala
            20   25   30
Gln Lys Ile Pro Asp Trp Ala Ser Ala Gly Ile Ser Ala Ile Trp Ile
        35   40   45
Pro Pro Ala Ser Lys Gly Met Gly Gly Ala Tyr Ser Met Gly Tyr Asp
    50   55   60
Pro Tyr Asp Phe Phe Asp Leu Gly Glu Tyr Asp Gln Lys Gly Thr Val
65   70   75   80
Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Val Asn Met Ile Asn Thr
                85   90   95
Ala His Ala Tyr Gly Met Lys Val Ile Ala Asp Ile Val Ile Asn His
            100   105   110
Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Asn Asp Tyr Thr
        115   120   125
Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr
    130   135   140
Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe
145   150   155   160
Gly Gly Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp
                165   170   175
Leu Trp Ala Ser Gln Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly
            180   185   190
Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Ala Pro Trp Val
        195   200   205
Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr
    210   215   220
Trp Asp Thr Asn Val Asp Ala Val Leu Asn Trp Ala Tyr Ser Ser Gly
225   230   235   240
Ala Lys Val Phe Asp Phe Ala Leu Tyr Tyr Lys Met Asp Glu Ala Phe
                245   250   255
Asp Asn Asn Asn Ile Pro Ala Leu Val Asp Ala Leu Arg Tyr Gly Gln
            260   265   270
Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn
        275   280   285
His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile
    290   295   300
Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu
305   310   315   320
Trp Leu Asn Lys Asp Thr Leu Lys Asn Leu Ile Trp Ile His Asp Asn
                325   330   335
Leu Ala Gly Gly Ser Thr Ser Ile Val Tyr Tyr Asp Ser Asp Glu Met
            340   345   350
Ile Phe Val Arg Asn Gly Tyr Gly Ser Lys Pro Gly Leu Ile Thr Tyr
        355   360   365
Ile Asn Leu Gly Ser Ser Lys Val Gly Arg Trp Val Tyr Val Pro Lys
    370   375   380
Phe Ala Gly Ala Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp
385   390   395   400
Val Asp Lys Trp Val Asp Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro
                405   410   415
Ala His Asp Pro Ala Asn Gly Tyr Tyr Gly Tyr Ser Val Trp Ser Tyr
            420   425   430
Cys Gly Val Gly
        435

<210> 13
<211> 1311
<212> DNA

<213> Artificial Sequence

<220>
<223> Synthetically generated oligonucleotide

<400> 13
atggccaagt acctggagct cgaagagggc ggggtcataa tgcaggcgtt ctactgggac 60
gtgccttcag gaggaatatg gtgggacaca atacggcaga agataccgga gtggtacgat 120
gccggaatct ccgcaatatg gattcccccg gcgagcaagg gcatgggcgg cgcctattcg 180
atgggctacg acccctacga cttctttgac ctcggtgagt atgaccagaa gggaacggta 240
gagacgcgct ttggctccaa gcaggagctc gtgaacatga taaacacggc acatgcctac 300
ggcataaagg tcatagcgga catcgtcata aaccaccgcg caggcggaga cctcgagtgg 360
aacccgttcg ttggggacta cacctggacg gacttctcaa aggtggcctc gggcaaatat 420
actgccaact acctcgactt ccaccccaac gaggtcaagt gctgtgacga gggcacattt 480
ggaggcttcc cagacatagc ccacgagaag agctgggacc agcactggct ctgggcgagc 540
gatgagagct acgccgccta cctaaggagc atcggcgttg atgcctggcg cttcgactac 600
gtcaagggct acggagcgtg ggtcgtcaag gactggctgg actggtgggg aggctgggcc 660
gtcggggagt actgggacac aaacgttgat gcactgctca actgggccta ctcgagcgat 720
gcaaaagtct tcgacttccc gctctactac aagatggatg aggcctttga caacaaaaac 780
attccagcgc tcgtctctgc ccttcagaac ggccagactg ttgtctcccg cgacccgttc 840
aaggccgtaa cctttgtagc aaaccacgac accgatataa tctggaacaa gtatccagcc 900
tacgcgttca tcctcaccta cgagggccag ccgacaatat tctaccgcga ctacgaggag 960
tggctcaaca aggataagct caagaacctc atctggatac atgacaacct cgccggagga 1020
agcactgaca tagtctacta cgataacgat gaactcatct tcgtcaggaa cggctacggg 1080
gacaagccgg ggcttataac ctacatcaac ctaggctcga gcaaggccgg aaggtgggtt 1140
tatgtgccga agttcgcggg cgcgtgcatc cacgagtata ctggtaacct cggaggctgg 1200
gtagacaagt acgtctactc aagcggctgg gtctatctcg aagctccagc ttacgaccct 1260
gccaacgggc agtatggcta ctccgtgtgg agctactgcg gtgttggctg a 1311

<210> 14
<211> 436
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetically generated peptide

<400> 14
Met Ala Lys Tyr Leu Glu Leu Glu Glu Gly Gly Val Ile Met Gln Ala
 1   5   10   15
Phe Tyr Trp Asp Val Pro Ser Gly Gly Ile Trp Trp Asp Thr Ile Arg
            20   25   30
Gln Lys Ile Pro Glu Trp Tyr Asp Ala Gly Ile Ser Ala Ile Trp Ile
        35   40   45
Pro Pro Ala Ser Lys Gly Met Gly Gly Ala Tyr Ser Met Gly Tyr Asp
    50   55   60
Pro Tyr Asp Phe Phe Asp Leu Gly Glu Tyr Asp Gln Lys Gly Thr Val
65   70   75   80
Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Val Asn Met Ile Asn Thr
                85   90   95
Ala His Ala Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His
            100   105   110
Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Gly Asp Tyr Thr
        115   120   125
Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr
    130   135   140
Leu Asp Phe His Pro Asn Glu Val Lys Cys Cys Asp Glu Gly Thr Phe
145   150   155   160
Gly Gly Phe Pro Asp Ile Ala His Glu Lys Ser Trp Asp Gln His Trp
                165   170   175
Leu Trp Ala Ser Asp Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly
            180   185   190
Val Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp Val
        195   200   205
Val Lys Asp Trp Leu Asp Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr
    210   215   220
Trp Asp Thr Asn Val Asp Ala Leu Leu Asn Trp Ala Tyr Ser Ser Asp
225   230   235   240

```
Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Glu Ala Phe
                245                 250                 255
Asp Asn Lys Asn Ile Pro Ala Leu Val Ser Ala Leu Gln Asn Gly Gln
            260                 265                 270
Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn
        275                 280                 285
His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile
    290                 295                 300
Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu
305                 310                 315                 320
Trp Leu Asn Lys Asp Lys Leu Lys Asn Leu Ile Trp Ile His Asp Asn
                325                 330                 335
Leu Ala Gly Gly Ser Thr Asp Ile Val Tyr Tyr Asp Asn Asp Glu Leu
            340                 345                 350
Ile Phe Val Arg Asn Gly Tyr Gly Asp Lys Pro Gly Leu Ile Thr Tyr
        355                 360                 365
Ile Asn Leu Gly Ser Ser Lys Ala Gly Arg Trp Val Tyr Val Pro Lys
    370                 375                 380
Phe Ala Gly Ala Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp
385                 390                 395                 400
Val Asp Lys Tyr Val Tyr Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro
                405                 410                 415
Ala Tyr Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr
            420                 425                 430
Cys Gly Val Gly
        435
```

<210> 15
<211> 1311
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetically generated oligonucleotide

<400> 15

```
atggccaagt actccgagct ggaagagggc gggctcataa tgcaggcctt ctactgggac      60
gtccccatgg gaggaatctg gtgggacacg atagcccaga agatacccga ctgggcaagc     120
gccgggattt cggcgatatg gattcccccg gcgagcaagg gcatgggcgg cgcctattcg     180
atgggctacg acccctacga cttctttgac ctcggtgagt acgaccagaa gggaacggta     240
gagacgcgct ttggctccaa gcaggagctc gtgaacatga taaacacggc ccatgcctac     300
ggcataaagg tcatagcgga catcgtcata aaccaccgcg caggcggaga cctcgagtgg     360
aacccgttcg ttggggacta cacctggacg gacttctcaa aggtggcctc gggcaaatat     420
actgccaact acctcgactt ccacccgaac gagctccatg cgggcgattc cggaacattt     480
ggaggctatc ccgacatatg ccacgacaag agctgggacc agtactggct ctgggccagc     540
caggagagct acgcggcata tctcaggagc atcggcatcg atgcctggcg cttcgactac     600
gtcaagggct acggagcgtg ggtcgtcaag gactggctgg actggtgggg aggctgggcc     660
gtcggggagt actgggacac aaacgttgat gcactgctca actgggccta ctcgagcgat     720
gcaaaagtct tcgacttccc gctctactac aagatggatg aggcctttga caacaaaaac     780
attccagcgc tcgtctctgc ccttcagaac ggccagactg ttgtctcccg cgacccgttc     840
aaggccgtaa cctttgtagc aaaccacgac accgatataa tttggaacaa gtacccggcc     900
tacgccttca tcctcaccta cgagggccag ccgacgatat tctaccgcga ctacgaggag     960
tggctcaaca aggacaggct caagaacctc atctggatac acgaccacct tgccggtgga    1020
agcactgaca tcgtttacta cgacaacgac gagctgatat tcgtgagaaa cggctacgga    1080
agcaagccgg gactgataac atacatcaac ctcgcctcaa gcaaagccgg aaggtgggtt    1140
tatgtgccga agttcgcggg cgcgtgcatc cacgagtata ctggtaacct cggaggctgg    1200
gtagacaagt acgtctactc aagcggctgg gtctatctcg aagctccagc ttacgaccct    1260
gccaacgggc agtatggcta ctccgtgtgg agctattgcg gtgttgggtg a             1311
```

<210> 16
<211> 436
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetically generated peptide

<400> 16

```
Met Ala Lys Tyr Ser Glu Leu Glu Glu Gly Gly Leu Ile Met Gln Ala
 1               5                  10                  15
Phe Tyr Trp Asp Val Pro Met Gly Gly Ile Trp Trp Asp Thr Ile Ala
            20                  25                  30
Gln Lys Ile Pro Asp Trp Ala Ser Ala Gly Ile Ser Ala Ile Trp Ile
        35                  40                  45
Pro Pro Ala Ser Lys Gly Met Gly Gly Ala Tyr Ser Met Gly Tyr Asp
    50                  55                  60
Pro Tyr Asp Phe Phe Asp Leu Gly Glu Tyr Asp Gln Lys Gly Thr Val
65                  70                  75                  80
Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Val Asn Met Ile Asn Thr
                85                  90                  95
Ala His Ala Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His
            100                 105                 110
Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Gly Asp Tyr Thr
        115                 120                 125
Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr
    130                 135                 140
Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe
145                 150                 155                 160
Gly Gly Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp
                165                 170                 175
Leu Trp Ala Ser Gln Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly
            180                 185                 190
Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp Val
        195                 200                 205
Val Lys Asp Trp Leu Asp Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr
    210                 215                 220
Trp Asp Thr Asn Val Asp Ala Leu Leu Asn Trp Ala Tyr Ser Ser Asp
225                 230                 235                 240
Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Glu Ala Phe
                245                 250                 255
Asp Asn Lys Asn Ile Pro Ala Leu Val Ser Ala Leu Gln Asn Gly Gln
            260                 265                 270
Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn
        275                 280                 285
His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile
    290                 295                 300
Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu
305                 310                 315                 320
Trp Leu Asn Lys Asp Arg Leu Lys Asn Leu Ile Trp Ile His Asp His
                325                 330                 335
Leu Ala Gly Gly Ser Thr Asp Ile Val Tyr Tyr Asp Asn Asp Glu Leu
            340                 345                 350
Ile Phe Val Arg Asn Gly Tyr Gly Ser Lys Pro Gly Leu Ile Thr Tyr
        355                 360                 365
Ile Asn Leu Ala Ser Ser Lys Ala Gly Arg Trp Val Tyr Val Pro Lys
    370                 375                 380
Phe Ala Gly Ala Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp
385                 390                 395                 400
Val Asp Lys Tyr Val Tyr Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro
                405                 410                 415
Ala Tyr Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr
            420                 425                 430
Cys Gly Val Gly
        435
```

<210> 17
<211> 1311
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetically generated oligonucleotide

<400> 17
atggccaagt actccgagct ggaaggggc gggctcataa tgcaggcctt ctactgggac 60
gtccccatgg gaggaatctg gtgggacacg atagcccaga agatacccga ctgggcaagc 120

```
gccgggattt cggcgatatg gattcccccg gcgagcaagg gcatgggcgg cgcctattcg     180
atgggctacg acccctacga cttctttgac ctcggtgagt acgaccagga gggaacggta     240
gagacgcgct ttggctccaa gcaggagctc gtgaacatga taaacacggc ccatgcctac     300
ggcataaagg tcatagcgga catcgtcata aaccaccgcg caggcggaga cctcgagtgg     360
aacccgttcg ttggggacta cacctggacg gacttctcaa aggtggcctc gggcaaatat     420
actgccaact acctcgactt ccaccccaac gaggtcaagt gctgtgacga gggcacattt     480
ggaggcttcc cagacatagc ccacgagaag agctgggacc agcactggct ctgggcgagc     540
gatgagagct acgccgccta cctaaggagc atcggcgttg atgcctggcg cttcgactac     600
gtcaagggct acggagcgtg ggtcgtcaag gactggctgg actggtgggg aggctgggcc     660
gtcggggagt actgggacac aaacgttgat gcactgctca actgggccta ctcgagcgat     720
gcaaaagtct tcgacttccc gctctactac aagatggacg cggcctttga caacaagaac     780
attcccgcac tcgtcgaggc cctcaagaac gggggcacag tcgtcagccg cgacccgttt     840
aaggccgtaa ccttcgttgc aaaccacgac accgatataa tctggaacaa gtatccagcc     900
tacgcgttca tcctcaccta cgagggccag ccgacaatat tctaccgcga ctacgaggag     960
tggctcaaca aggataagct caagaacctc atctggatac atgacaacct cgccggagga    1020
agcacgagca tagtttacta cgacagcgac gagatgatct tcgtgaggaa cggctatgga    1080
agcaagcctg gccttataac ttacatcaac ctcggctcga gcaaggttgg aaggtgggtt    1140
tacgttccga agttcgcagg ctcgtgcata cacgagtaca ccggcaatct cggcggctgg    1200
gtggacaagt gggtggactc aagcggctgg gtctacctcg aggctcctgc ccacgacccg    1260
gccaacggcc agtacggcta ctccgtctgg agctactgcg gtgttgggtg a              1311

<210> 18
<211> 436
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetically generated peptide

<400> 18
Met Ala Lys Tyr Ser Glu Leu Glu Gly Gly Gly Leu Ile Met Gln Ala
 1               5                  10                  15
Phe Tyr Trp Asp Val Pro Met Gly Gly Ile Trp Trp Asp Thr Ile Ala
            20                  25                  30
Gln Lys Ile Pro Asp Trp Ala Ser Ala Gly Ile Ser Ala Ile Trp Ile
        35                  40                  45
Pro Pro Ala Ser Lys Gly Met Gly Gly Ala Tyr Ser Met Gly Tyr Asp
    50                  55                  60
Pro Tyr Asp Phe Phe Asp Leu Gly Glu Tyr Asp Gln Glu Gly Thr Val
65                  70                  75                  80
Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Val Asn Met Ile Asn Thr
                85                  90                  95
Ala His Ala Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His
            100                 105                 110
Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Gly Asp Tyr Thr
        115                 120                 125
Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr
    130                 135                 140
Leu Asp Phe His Pro Asn Glu Val Lys Cys Cys Asp Glu Gly Thr Phe
145                 150                 155                 160
Gly Gly Phe Pro Asp Ile Ala His Glu Lys Ser Trp Asp Gln His Trp
                165                 170                 175
Leu Trp Ala Ser Asp Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly
            180                 185                 190
Val Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp Val
        195                 200                 205
Val Lys Asp Trp Leu Asp Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr
    210                 215                 220
Trp Asp Thr Asn Val Asp Ala Leu Leu Asn Trp Ala Tyr Ser Ser Asp
225                 230                 235                 240
Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Ala Ala Phe
                245                 250                 255
Asp Asn Lys Asn Ile Pro Ala Leu Val Glu Ala Leu Lys Asn Gly Gly
            260                 265                 270
Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn
        275                 280                 285
His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile
    290                 295                 300
```

Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu
305 310 315 320
Trp Leu Asn Lys Asp Lys Leu Lys Asn Leu Ile Trp Ile His Asp Asn
325 330 335
Leu Ala Gly Gly Ser Thr Ser Ile Val Tyr Tyr Asp Ser Asp Glu Met
340 345 350
Ile Phe Val Arg Asn Gly Tyr Gly Ser Lys Pro Gly Leu Ile Thr Tyr
355 360 365
Ile Asn Leu Gly Ser Ser Lys Val Gly Arg Trp Val Tyr Val Pro Lys
370 375 380
Phe Ala Gly Ser Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp
385 390 395 400
Val Asp Lys Trp Val Asp Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro
405 410 415
Ala His Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr
420 425 430
Cys Gly Val Gly
435

<210> 19
<211> 1311
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetically generated oligonucleotide

<400> 19
atggccaagt acctggagct cgaagagggc gggctcataa tgcaggcctt ctactgggac 60
gtccccatgg gaggaatctg gtgggacacg atagcccaga agatacccga ctgggcaagc 120
gccgggattt cggcgatatg gattcctccc gcgagcaagg gtatgagcgg cggctattcg 180
atgggctacg acccctacga ttattttgac cttggtgagt actaccagaa gggaacggtg 240
gaaacgaggt tcggctcaaa gcaggagctc ataaacatga taaacacggc ccatgcctac 300
ggcataaagg tcatagcgga catcgtcata aaccaccgcg caggcggaga cctcgagtgg 360
aacccgttcg ttggggacta cacctggacg gacttctcaa aggtggcctc gggcaaatat 420
actgccaact acctcgactt ccacccgaac gagctccatg cgggcgattc cggaacattt 480
ggaggctatc ccgacatatg ccacgacaag agctgggacc agtactggct ctgggccagc 540
caggagagct acgcggcata tctcaggagc atcggcatcg atgcctggcg cttcgactac 600
gtcaagggct atgctccctg ggtcgtcaag gactggctga actggtgggg gggctgggcg 660
gttggagagt actgggacac caacgtcgac gctgttctca actgggcata ctcgagcggt 720
gccaaggtct ttgacttcgc cctctactac aagatggatg aggcctttga caacaaaaac 780
attccagcgc tcgtctctgc ccttcagaac ggccagactg ttgtctcccg cgacccgttc 840
aaggccgtaa cctttgtagc aaaccacgac accgatataa tttggaacaa gtacccggcc 900
tacgccttca tcctcaccta cgagggccag ccgacgatat tctaccgcga ctacgaggag 960
tggctcaaca aggacaggct caagaacctc atctggatac acgaccacct cgccggtgga 1020
agcactgaca tcgtttacta cgacaacgac gagctgatat tcgtgagaaa cggctacgga 1080
agcaagccgg gactgataac atacatcaac ctcgcctcaa gcaaagccgg aaggtgggtt 1140
tatgtgccga agttcgcggg cgcgtgcatc cacgagcata ctggtaacct cggaggctgg 1200
gtagacaagt acgtctactc aagcggctgg gtctatctcg aagctccagc ttacgaccct 1260
gccaacgggc agtatggcta ctccgtgtgg agctactgcg gtgttggctg a 1311

<210> 20
<211> 436
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetically generated peptide

<400> 20
Met Ala Lys Tyr Leu Glu Leu Glu Glu Gly Gly Leu Ile Met Gln Ala
1 5 10 15
Phe Tyr Trp Asp Val Pro Met Gly Gly Ile Trp Trp Asp Thr Ile Ala
20 25 30
Gln Lys Ile Pro Asp Trp Ala Ser Ala Gly Ile Ser Ala Ile Trp Ile
35 40 45
Pro Pro Ala Ser Lys Gly Met Ser Gly Gly Tyr Ser Met Gly Tyr Asp
50 55 60

```
Pro Tyr Asp Tyr Phe Asp Leu Gly Glu Tyr Tyr Gln Lys Gly Thr Val
65                  70                  75                  80
Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Ile Asn Met Ile Asn Thr
                85                  90                  95
Ala His Ala Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His
            100                 105                 110
Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Gly Asp Tyr Thr
        115                 120                 125
Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr
    130                 135                 140
Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe
145                 150                 155                 160
Gly Gly Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp
                165                 170                 175
Leu Trp Ala Ser Gln Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly
            180                 185                 190
Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Ala Pro Trp Val
        195                 200                 205
Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr
    210                 215                 220
Trp Asp Thr Asn Val Asp Ala Val Leu Asn Trp Ala Tyr Ser Ser Gly
225                 230                 235                 240
Ala Lys Val Phe Asp Phe Ala Leu Tyr Tyr Lys Met Asp Glu Ala Phe
                245                 250                 255
Asp Asn Lys Asn Ile Pro Ala Leu Val Ser Ala Leu Gln Asn Gly Gln
            260                 265                 270
Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn
        275                 280                 285
His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile
    290                 295                 300
Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu
305                 310                 315                 320
Trp Leu Asn Lys Asp Arg Leu Lys Asn Leu Ile Trp Ile His Asp His
                325                 330                 335
Leu Ala Gly Gly Ser Thr Asp Ile Val Tyr Tyr Asp Asn Asp Glu Leu
            340                 345                 350
Ile Phe Val Arg Asn Gly Tyr Gly Ser Lys Pro Gly Leu Ile Thr Tyr
        355                 360                 365
Ile Asn Leu Ala Ser Ser Lys Ala Gly Arg Trp Val Tyr Val Pro Lys
    370                 375                 380
Phe Ala Gly Ala Cys Ile His Glu His Thr Gly Asn Leu Gly Gly Trp
385                 390                 395                 400
Val Asp Lys Tyr Val Tyr Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro
                405                 410                 415
Ala Tyr Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr
            420                 425                 430
Cys Gly Val Gly
        435
```

<210> 21
<211> 1311
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetically generated oligonucleotide

<400> 21

```
atggccaagt actccgagct ggaagagggc ggcgttataa tgcaggcctt ctactgggac      60
gtcccaggtg gaggaatctg gtgggacacc atcaggagca agataccgga gtggtacgag     120
gcgggaatat ccgccatttg gattcctccc gggagcaagg gtatgagcgg cggctattcg     180
atgggctacg acccctacga tgatttggac ctgggtgagt actaccagaa gggaacggtg     240
gaaacgaggt tcggctcaaa gcaggagctc ataaacatga taaacacggc ccatgcctac     300
ggcataaagg tcatagcgga catcgtcata aaccaccgcg caggcggaga cctcgagtgg     360
aacccgttcg ttggggacta cacctggacg gacttctcaa aggtggcctc gggcaaatat     420
actgccaact acctcgactt ccacccgaac gagctccatg cgggcgattc cggaacattt     480
ggaggctatc ccgacatatg ccacgacaag agctgggacc agtactggct ctgggccagc     540
caggagagct acgcggtata tctcaggagc atcggcatcg atgcctggcg cttcgactac     600
```

gtcaagggct acggagcgtg ggtcgtcaag gactggctgg actggtgggg aggctgggcc 660
gtcggggagt actgggacac aaacgttgat gcactgctca actgggccta ctcgagcgat 720
gcaaaagtct tcgacttccc gctctactac aagatggatg aggcctttga caacaaaaac 780
attccagcgc tcgtctctgc ccttcagaac ggccagactg ttgtctcccg cgacccgttc 840
aaggccgtaa cctttgtagc aaaccacgac accgatataa tttggaacaa gtacccggcc 900
tacgccttca tcctcaccta cgagggccag ccgacgatat tctaccgcga ctacgaggag 960
tggctcaaca aggacaggct caagaacctc atctggatac acgactacct cgccggtgga 1020
agcactgaca tcgtttacta cgacaacgac gagctgatat tcgtgagaaa cggctacgga 1080
agcaagccgg gactgataac atacatcaac ctcgcctcaa gcaaagccgg aaggtgggtt 1140
tatgtgccga agttcgcggg cgcgtgcatc cacgagtata ctggtaacct cggaggctgg 1200
gtagacaagt acgtctactc aagcggctgg gtctatctcg aagctccagc ttacgaccct 1260
gccaacgggc agtatggcta ctccgtgtgg agctattgcg gtgttggctg a 1311

<210> 22
<211> 436
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetically generated peptide

<400> 22
Met Ala Lys Tyr Ser Glu Leu Glu Glu Gly Gly Val Ile Met Gln Ala
1 5 10 15
Phe Tyr Trp Asp Val Pro Gly Gly Gly Ile Trp Trp Asp Thr Ile Arg
20 25 30
Ser Lys Ile Pro Glu Trp Tyr Glu Ala Gly Ile Ser Ala Ile Trp Ile
35 40 45
Pro Pro Gly Ser Lys Gly Met Ser Gly Gly Tyr Ser Met Gly Tyr Asp
50 55 60
Pro Tyr Asp Asp Leu Asp Leu Gly Glu Tyr Tyr Gln Lys Gly Thr Val
65 70 75 80
Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Ile Asn Met Ile Asn Thr
85 90 95
Ala His Ala Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His
100 105 110
Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Gly Asp Tyr Thr
115 120 125
Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr
130 135 140
Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe
145 150 155 160
Gly Gly Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp
165 170 175
Leu Trp Ala Ser Gln Glu Ser Tyr Ala Val Tyr Leu Arg Ser Ile Gly
180 185 190
Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp Val
195 200 205
Val Lys Asp Trp Leu Asp Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr
210 215 220
Trp Asp Thr Asn Val Asp Ala Leu Leu Asn Trp Ala Tyr Ser Ser Asp
225 230 235 240
Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Glu Ala Phe
245 250 255
Asp Asn Lys Asn Ile Pro Ala Leu Val Ser Ala Leu Gln Asn Gly Gln
260 265 270
Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn
275 280 285
His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile
290 295 300
Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu
305 310 315 320
Trp Leu Asn Lys Asp Arg Leu Lys Asn Leu Ile Trp Ile His Asp Tyr
325 330 335
Leu Ala Gly Gly Ser Thr Asp Ile Val Tyr Tyr Asp Asn Asp Glu Leu
340 345 350
Ile Phe Val Arg Asn Gly Tyr Gly Ser Lys Pro Gly Leu Ile Thr Tyr
355 360 365

Ile Asn Leu Ala Ser Ser Lys Ala Gly Arg Trp Val Tyr Val Pro Lys
370 375 380
Phe Ala Gly Ala Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp
385 390 395 400
Val Asp Lys Tyr Val Tyr Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro
405 410 415
Ala Tyr Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr
420 425 430
Cys Gly Val Gly
435

<210> 23
<211> 1311
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetically generated oligonucleotide

<400> 23
atggccaagt actccgagct ggaagagggc ggcgttatag tgcaggcctt ctactgggac 60
gtcccaggtg gaggaatctg gtgggacacc atcaggagca agataccgga gtggtacgag 120
gcgggaatat ccgccatttg gattcccccg gcgagcaagg gcatgggcgg cgcctattcg 180
atgggctacg acccctacga cttctttgac ctcggtgagt acgaccagaa gggaacggta 240
gagacgcgct ttggctccaa gcaggagctc gtgaacatga taaacacggc ccatgcctac 300
ggcataaagg tcatagcgga catcgtcata aaccaccgcg caggcggaga cctcgagtgg 360
aacccgttcg ttggggacta cacctggacg gacttctcaa aggtggcctc gggcaaatat 420
actgccaact acctcgactt ccacccgaac gagctccatg cgggcgattc cggaacattt 480
ggaggctatc ccgacatatg ccacgacaag agctgggacc agtactggct ctgggccagc 540
caggagagct acgcggcata tctcaggagc atcggcatcg atgcctggcg cttcgactac 600
gtcaagggct acggagcgtg ggtcgtcaag gactggctgg actggtgggg aggctgggcc 660
gtcggggagt actgggacac aaacgttgat gcactgctca actgggccta ctcgagcgat 720
gcaaaagtct tcgacttccc gctctactac aagatggatg aggcctttga caacaaaaac 780
attccagcgc tcgtctctgc ccttcagaac ggccagactg ttgtctcccg cgacccgttc 840
aaggccgtaa cctttgtagc aaaccacgac accgatataa tctggaacaa gtatccagcc 900
tacgcgttca tcctcaccta cgagggccag ccgacaatat tctaccgcga ctacgaggag 960
tggctcaaca aggataagct caagaacctc atctggatac atgacaacct cgccggagga 1020
agcatgagca tagtttacta cgacagcgac gagatgatct tcgtgaggaa cggctatgga 1080
agcaagcctg gccttataac ttacatcaac ctcggctcga gcaaggttgg aaggtgggtc 1140
tacgttccga agttcgcggg agcgtgcatc cacgagtaca ccggcaacct cggcggctgg 1200
gtggacaagt gggtggactc aagcgggtgg gtgtacctcg aggcccctgc ccacgacccg 1260
gccaacggct attacggcta ctccgtctgg agctattgcg gtgttggctg a 1311

<210> 24
<211> 436
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetically generated peptide

<400> 24
Met Ala Lys Tyr Ser Glu Leu Glu Glu Gly Gly Val Ile Val Gln Ala
1 5 10 15
Phe Tyr Trp Asp Val Pro Gly Gly Gly Ile Trp Trp Asp Thr Ile Arg
20 25 30
Ser Lys Ile Pro Glu Trp Tyr Glu Ala Gly Ile Ser Ala Ile Trp Ile
35 40 45
Pro Pro Ala Ser Lys Gly Met Gly Gly Ala Tyr Ser Met Gly Tyr Asp
50 55 60
Pro Tyr Asp Phe Phe Asp Leu Gly Glu Tyr Asp Gln Lys Gly Thr Val
65 70 75 80
Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Val Asn Met Ile Asn Thr
85 90 95
Ala His Ala Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His
100 105 110
Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Gly Asp Tyr Thr
115 120 125

```
Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr
    130                 135                 140
Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe
145                 150                 155                 160
Gly Gly Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp
                165                 170                 175
Leu Trp Ala Ser Gln Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly
            180                 185                 190
Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp Val
        195                 200                 205
Val Lys Asp Trp Leu Asp Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr
    210                 215                 220
Trp Asp Thr Asn Val Asp Ala Leu Leu Asn Trp Ala Tyr Ser Ser Asp
225                 230                 235                 240
Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Glu Ala Phe
                245                 250                 255
Asp Asn Lys Asn Ile Pro Ala Leu Val Ser Ala Leu Gln Asn Gly Gln
            260                 265                 270
Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn
        275                 280                 285
His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile
    290                 295                 300
Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu
305                 310                 315                 320
Trp Leu Asn Lys Asp Lys Leu Lys Asn Leu Ile Trp Ile His Asp Asn
                325                 330                 335
Leu Ala Gly Gly Ser Met Ser Ile Val Tyr Tyr Asp Ser Asp Glu Met
            340                 345                 350
Ile Phe Val Arg Asn Gly Tyr Gly Ser Lys Pro Gly Leu Ile Thr Tyr
        355                 360                 365
Ile Asn Leu Gly Ser Ser Lys Val Gly Arg Trp Val Tyr Val Pro Lys
    370                 375                 380
Phe Ala Gly Ala Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp
385                 390                 395                 400
Val Asp Lys Trp Val Asp Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro
                405                 410                 415
Ala His Asp Pro Ala Asn Gly Tyr Tyr Gly Tyr Ser Val Trp Ser Tyr
            420                 425                 430
Cys Gly Val Gly
        435
```

<210> 25
<211> 1311
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetically generated oligonucleotide

<400> 25
```
atggccaagt acctggagct cgaagagggc gggctcataa tgcaggcctt ctactgggac      60
gtccccatgg gaggaatctg gtgggacacg atagcccaga agatacccga ctgggcaagc     120
gccgggattt cggcgatatg gattcctccc gcgagcaagg gtatgagcgg cggctattcg     180
atgggctacg acccctacga ttattttgac ctcggtgagt actaccagaa gggaacggtg     240
gaaacgaggt tcggctcaaa gcaggagctc ataaacatga taaacaccgc ccacgcctat     300
ggcatgaagg taatagccga tatagtcatc aaccaccgcg ccggcggtga cctggagtgg     360
aacccttcg tgaacgacta tacctggacc gacttctcaa aggtcgcgtc gggtaaatac     420
acggccaact acctcgactt ccacccgaac gagctccatg cgggcgattc cggaacattt     480
ggaggctatc ccgacatatg ccacgacaag agctgggacc agtactggct ctgggccagc     540
caggagagct acgcggcata tctcaggagc atcggcatcg atgcctggcg cttcgactac     600
gtcaagggct atgctccctg ggtcgtcaag gactggctga actggtgggg aggctgggcg     660
gttggagagt actgggacac caacgtcgac gctgttctca actgggcata ctcgagcggt     720
gccaaggtct ttgacttcgc cctctactac aagatggacg aggccttcga taacaacaac     780
attcccgccc tggtgggcgc cctcagatac ggtcagacag tggtcagccg cgacccgttc     840
aaggctgtga cgtttgtagc caaccacgat accgatataa tctggaacaa gtatccagcc     900
tacgcgttca tcctcaccta cgagggccag ccgacaatat tctaccgcga ctacgaggag     960
tggctcaaca aggataagct caagaacctc atctggatac atgacaacct cgccggagga    1020
agcaccgaca tagtctacta cgataacgat gaactcatct tcgtcaggca cggctacggg    1080
```

gacaagccgg ggcttataac ctacatcaac ctaggctcga gcaaggccgg aaggtgggtt 1140
tacgttccga agttcgcagg ctcgtgcata cacgagtaca ccggcaatct cggcggctgg 1200
gtggacaagt gggtggactc aagcggctgg gtctacctcg aggctcctgc ccacgacccg 1260
gccaacggcc agtacggcta ctccgtctgg agctattgcg gtgttgggtg a 1311

<210> 26
<211> 436
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetically generated peptide

<400> 26
Met Ala Lys Tyr Leu Glu Leu Glu Glu Gly Gly Leu Ile Met Gln Ala
 1   5   10   15
Phe Tyr Trp Asp Val Pro Met Gly Gly Ile Trp Trp Asp Thr Ile Ala
            20   25   30
Gln Lys Ile Pro Asp Trp Ala Ser Ala Gly Ile Ser Ala Ile Trp Ile
        35   40   45
Pro Pro Ala Ser Lys Gly Met Ser Gly Gly Tyr Ser Met Gly Tyr Asp
    50   55   60
Pro Tyr Asp Tyr Phe Asp Leu Gly Glu Tyr Tyr Gln Lys Gly Thr Val
65   70   75   80
Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Ile Asn Met Ile Asn Thr
                85   90   95
Ala His Ala Tyr Gly Met Lys Val Ile Ala Asp Ile Val Ile Asn His
            100   105   110
Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Asn Asp Tyr Thr
        115   120   125
Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr
    130   135   140
Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe
145   150   155   160
Gly Gly Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp
                165   170   175
Leu Trp Ala Ser Gln Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly
            180   185   190
Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Ala Pro Trp Val
        195   200   205
Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr
    210   215   220
Trp Asp Thr Asn Val Asp Ala Val Leu Asn Trp Ala Tyr Ser Ser Gly
225   230   235   240
Ala Lys Val Phe Asp Phe Ala Leu Tyr Tyr Lys Met Asp Glu Ala Phe
                245   250   255
Asp Asn Asn Asn Ile Pro Ala Leu Val Gly Ala Leu Arg Tyr Gly Gln
            260   265   270
Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn
        275   280   285
His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile
    290   295   300
Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu
305   310   315   320
Trp Leu Asn Lys Asp Lys Leu Lys Asn Leu Ile Trp Ile His Asp Asn
                325   330   335
Leu Ala Gly Gly Ser Thr Asp Ile Val Tyr Tyr Asp Asn Asp Glu Leu
            340   345   350
Ile Phe Val Arg His Gly Tyr Gly Asp Lys Pro Gly Leu Ile Thr Tyr
        355   360   365
Ile Asn Leu Gly Ser Ser Lys Ala Gly Arg Trp Val Tyr Val Pro Lys
    370   375   380
Phe Ala Gly Ser Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp
385   390   395   400
Val Asp Lys Trp Val Asp Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro
                405   410   415
Ala His Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr
            420   425   430

Cys Gly Val Gly
435

<210> 27
<211> 1311
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetically generated oligonucleotide

<400> 27
atggcaaagt attccgagct cgaagagggc ggcgttataa tgcaggcctt ctactgggac 60
gtcccaggtg gaggaatctg gtgggacacc atcaggagca agataccgga gtggtacgag 120
gcgggaatat ccgccatttg gattcctccc gcgagcaagg gtatgagcgg cggctattcg 180
atgggctacg acccctacga ttattttgac ctcggtgagt actaccagaa gggaacggtg 240
gaaacgaggt tcggctcaaa gcaggagctc ataaacatga taaacacggc ccatgcctac 300
ggcataaagg tcatagcgga catcgtcata aaccaccgcg caggcggaga cctcgagtgg 360
aacccgttcg ttggggacta cacctggacg gacttctcaa aggtggcctc gggcaaatat 420
actgccaact acctcgactt ccacccgaac gagctccatg cgggcgattc cggaacattt 480
ggaggctatc ccgacatatg ccacgacaag agctgggacc agtactggct ctgggccagc 540
caggagagct acgcggcata tctcaggagc atcggcatcg atgcctggcg cttcgactac 600
gtcaagggct atgctccctg ggtcgtcaag gactggctga actggtgggg aggctgggcg 660
gttggagagt actgggacac caacgtcgac gctgttctca actgggcata ctcgagcggt 720
gccaaggtct ttgacttcgc cctctactac aagatggacg cggcctttga caacaagaac 780
attcccgcac tcgtcgaggc cctcaagaac gggggcacag tcgtcagccg cgacccgttt 840
aaggccgtaa ccttcgttgc aaaccacgac accgatataa tctggaacaa gtatccagcc 900
tacgcgttca tcctcaccta cgagggccag ccgacaatat tctaccgcga ctacgaggag 960
tggctcaaca aggataagct caagaacctc atctggatac atgacaacct cgccggagga 1020
agcactgaca tcgtttacta cgacaacgac gagctgatat tcgtgagaaa cggctacgga 1080
agcaagccgg gactgataac atacatcaac ctcgcgtcaa gcaaagccgg aaggtgggtt 1140
tacgttccga agttcgcagg ctcgtgcata cacgagtaca ccggcaatct cggcggctgg 1200
gtggacaagt gggtggactc aagcggctgg gtctacctcg aggctcctgc ccacgacccg 1260
gccaacggcc agtacggcta ctccgtctgg agctactgcg gtgttgggtg a 1311

<210> 28
<211> 436
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetically generated peptide

<400> 28
Met Ala Lys Tyr Ser Glu Leu Glu Glu Gly Gly Val Ile Met Gln Ala
1 5 10 15
Phe Tyr Trp Asp Val Pro Gly Gly Gly Ile Trp Trp Asp Thr Ile Arg
20 25 30
Ser Lys Ile Pro Glu Trp Tyr Glu Ala Gly Ile Ser Ala Ile Trp Ile
35 40 45
Pro Pro Ala Ser Lys Gly Met Ser Gly Gly Tyr Ser Met Gly Tyr Asp
50 55 60
Pro Tyr Asp Tyr Phe Asp Leu Gly Glu Tyr Tyr Gln Lys Gly Thr Val
65 70 75 80
Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Ile Asn Met Ile Asn Thr
85 90 95
Ala His Ala Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His
100 105 110
Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Gly Asp Tyr Thr
115 120 125
Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr
130 135 140
Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe
145 150 155 160
Gly Gly Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp
165 170 175
Leu Trp Ala Ser Gln Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly
180 185 190

```
Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Ala Pro Trp Val
        195                 200                 205
Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr
    210                 215                 220
Trp Asp Thr Asn Val Asp Ala Val Leu Asn Trp Ala Tyr Ser Ser Gly
225                 230                 235                 240
Ala Lys Val Phe Asp Phe Ala Leu Tyr Tyr Lys Met Asp Ala Ala Phe
                245                 250                 255
Asp Asn Lys Asn Ile Pro Ala Leu Val Glu Ala Leu Lys Asn Gly Gly
            260                 265                 270
Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn
        275                 280                 285
His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile
    290                 295                 300
Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu
305                 310                 315                 320
Trp Leu Asn Lys Asp Lys Leu Lys Asn Leu Ile Trp Ile His Asp Asn
                325                 330                 335
Leu Ala Gly Gly Ser Thr Asp Ile Val Tyr Tyr Asp Asn Asp Glu Leu
            340                 345                 350
Ile Phe Val Arg Asn Gly Tyr Gly Ser Lys Pro Gly Leu Ile Thr Tyr
        355                 360                 365
Ile Asn Leu Ala Ser Ser Lys Ala Gly Arg Trp Val Tyr Val Pro Lys
    370                 375                 380
Phe Ala Gly Ser Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp
385                 390                 395                 400
Val Asp Lys Trp Val Asp Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro
                405                 410                 415
Ala His Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr
            420                 425                 430
Cys Gly Val Gly
        435
```

<210> 29
<211> 1311
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetically generated oligonucleotide

<400> 29

```
atggccaagt acctggagct cgaagagggc gggctcataa tgcaggcctt ctactgggac      60
gtccccatgg gaggaatctg gtgggacacg gtagcccaga agatacccga ctgggcaagc     120
gccgggattt cggcgatatg gattcccccg gcgagcaagg gcatgggcgg cgcctattcg     180
atgggctacg acccctacga cttctttgac ctcggtgagt acgaccagaa gggaacggta     240
gagacgcgct ttggctccaa gcaggagctc gtgaacatga taaacacggc ccatgcctac     300
ggcataaagg tcatagcgga catcgtcata aaccaccgcg caggcggaga cctcgagtgg     360
aacccgttcg ttggggacta cacctggacg gacttctcaa aggtggtctc gggcaaatat     420
actgccaact acctcgactt ccacccgaac gagctccatg cgggcgattc cggaacattt     480
ggaggctatc ccgacatatg ccacgacaag agctgggacc agtactggct ctgggccagc     540
caggagagct acgcggcata tctcaggagc atcggcatcg atgcctggcg cttcgactac     600
gtcaagggct atgctccctg ggtcgtcaag gactggctga actggtgggg aggctgggcg     660
gttggagagt actgggacac caacgtcgac gctgttctca actgggcata ctcgagcggt     720
gccaaggtct ttgacttcgc cctctactac aagatggatg aggcctttga caacaaaaac     780
attccagcgc tcgtctctgc ccttcagaac ggccagactg ttgtctcccg cgacccgttc     840
aaggccgtaa cctttgtagc aaaccacgac accgatataa tctggaacaa gtaccttgct     900
tatgccttca tcctcaccta cgaaggccag cccgtcatat tctaccgcga ctacgaggag     960
tggctcaaca aggacaggtt gaacaacctc atatggatac acgaccacct cgcagggga    1020
agcaccgaca tagtctacta cgataacgat gaactcatct tcgtcaggaa cggctacggg    1080
gacaagccgg ggcttataac ctacatcaac ctaggctcga gcaaggccgg aaggtgggtt    1140
tatgtgccga agttcgcggg cgcgtgcatc cacgagtata ctggtaacct cggaggctgg    1200
gtagacaagt acgtctactc aagcggctgg gtctatctcg aagctccagc ttacgaccct    1260
gccaacgggc agtatggcta ctccgtgtgg agctactgcg gtgttgggtg a             1311
```

<210> 30
<211> 436
<212> PRT

<213> Artificial Sequence

<220>
<223> Synthetically generated peptide

<400> 30
Met Ala Lys Tyr Leu Glu Leu Glu Glu Gly Gly Leu Ile Met Gln Ala
 1   5   10   15
Phe Tyr Trp Asp Val Pro Met Gly Gly Ile Trp Trp Asp Thr Val Ala
            20   25   30
Gln Lys Ile Pro Asp Trp Ala Ser Ala Gly Ile Ser Ala Ile Trp Ile
        35   40   45
Pro Pro Ala Ser Lys Gly Met Gly Gly Ala Tyr Ser Met Gly Tyr Asp
    50   55   60
Pro Tyr Asp Phe Phe Asp Leu Gly Glu Tyr Asp Gln Lys Gly Thr Val
65   70   75   80
Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Val Asn Met Ile Asn Thr
                85   90   95
Ala His Ala Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His
            100   105   110
Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Gly Asp Tyr Thr
        115   120   125
Trp Thr Asp Phe Ser Lys Val Val Ser Gly Lys Tyr Thr Ala Asn Tyr
    130   135   140
Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe
145   150   155   160
Gly Gly Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp
                165   170   175
Leu Trp Ala Ser Gln Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly
            180   185   190
Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Ala Pro Trp Val
        195   200   205
Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr
    210   215   220
Trp Asp Thr Asn Val Asp Ala Val Leu Asn Trp Ala Tyr Ser Ser Gly
225   230   235   240
Ala Lys Val Phe Asp Phe Ala Leu Tyr Tyr Lys Met Asp Glu Ala Phe
                245   250   255
Asp Asn Lys Asn Ile Pro Ala Leu Val Ser Ala Leu Gln Asn Gly Gln
            260   265   270
Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn
        275   280   285
His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Leu Ala Tyr Ala Phe Ile
    290   295   300
Leu Thr Tyr Glu Gly Gln Pro Val Ile Phe Tyr Arg Asp Tyr Glu Glu
305   310   315   320
Trp Leu Asn Lys Asp Arg Leu Asn Asn Leu Ile Trp Ile His Asp His
                325   330   335
Leu Ala Gly Gly Ser Thr Asp Ile Val Tyr Tyr Asp Asn Asp Glu Leu
            340   345   350
Ile Phe Val Arg Asn Gly Tyr Gly Asp Lys Pro Gly Leu Ile Thr Tyr
        355   360   365
Ile Asn Leu Gly Ser Ser Lys Ala Gly Arg Trp Val Tyr Val Pro Lys
    370   375   380
Phe Ala Gly Ala Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp
385   390   395   400
Val Asp Lys Tyr Val Tyr Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro
                405   410   415
Ala Tyr Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr
            420   425   430
Cys Gly Val Gly
        435

<210> 31
<211> 1311
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetically generated oligonucleotide

<400> 31
atggcaaagt actccgagct ggaagagggc ggcgttataa tgcaggcctt ctactgggac 60
gtcccaggtg gaggaatctg gtgggacacc atcaggagca ggataccgga gtggtacgag 120
gcgggaatat ccgccatttg gattcccccg gcgagcaagg gcatgggcgg cgcctattcg 180
atgggctacg acccctacga cttctttgac ctcggtgagt acgaccagaa gggaacggta 240
gagacgcgct ttggctccaa gcaggagctc gtgaacatga taaacacggc ccatgcctac 300
ggcataaagg tcatagcgga catcgtcata aaccaccgcg caggcggaga cctcgagtgg 360
aacccgttcg ttggggacta cacctggacg gacttctcaa aggtggcctc gggcaaatat 420
actgccaact acctcgactt ccacccgaac gagctccatg cgggcgattc cggaacattt 480
ggaggctatc ccgacatatg ccacgacaag agctgggacc agtactggct ctgggccagc 540
caggagagct acgcggcata tctcaggagc atcggcatcg atgcctggcg ctttgactac 600
gtgaagggct acggagcgtg ggtcgtcaag gactggctca actggtgggg cggctgggcc 660
gttggcgagt actgggacac caacgttgat gcactcctca actgggccta ctcgagcggc 720
gccaaggtct tcgacttccc gctctactac aagatggacg aggccttcga taacaacaac 780
attcccgccc tggtggacgc cctcagatac ggtcagacag tggtcagccg cgacccgttc 840
aaggctgtga cgtttgtagc caaccacgat accgatataa tctggaacaa gtatccagcc 900
tacgcgttca tcctcaccta cgagggccag ccgacaatat tctaccgcga ctacgaggag 960
tggctcaaca aggataagct caagaacctc atctggatac atgacaacct ggccggagga 1020
agcacgagca tagtttacta cgacagcgac gagatgatct tcgtgaggac cggctatgga 1080
agcaagcctg gccttataac ttacatcaac ctcggctcga gcaaggttgg aaggtgggtt 1140
tatgtgccga agttcgcggg cgcgtgcatc cacgagtata ctggtaacct cggaggctgg 1200
gtagacaagt acgtctactc aagcggctgg gtctatctcg aagctccagc ttacgaccct 1260
gccaacgggc agtatggcta ctccgtgtgg agctattgcg gtgttggctg a 1311

<210> 32
<211> 436
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetically generated peptide

<400> 32
Met Ala Lys Tyr Ser Glu Leu Glu Glu Gly Gly Val Ile Met Gln Ala
 1   5   10   15
Phe Tyr Trp Asp Val Pro Gly Gly Gly Ile Trp Trp Asp Thr Ile Arg
            20   25   30
Ser Arg Ile Pro Glu Trp Tyr Glu Ala Gly Ile Ser Ala Ile Trp Ile
        35   40   45
Pro Pro Ala Ser Lys Gly Met Gly Gly Ala Tyr Ser Met Gly Tyr Asp
    50   55   60
Pro Tyr Asp Phe Phe Asp Leu Gly Glu Tyr Asp Gln Lys Gly Thr Val
65   70   75   80
Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Val Asn Met Ile Asn Thr
                85   90   95
Ala His Ala Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His
            100   105   110
Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Gly Asp Tyr Thr
        115   120   125
Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr
    130   135   140
Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe
145   150   155   160
Gly Gly Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp
                165   170   175
Leu Trp Ala Ser Gln Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly
            180   185   190
Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp Val
        195   200   205
Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr
    210   215   220
Trp Asp Thr Asn Val Asp Ala Leu Leu Asn Trp Ala Tyr Ser Ser Gly
225   230   235   240
Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Glu Ala Phe
                245   250   255

```
Asp Asn Asn Asn Ile Pro Ala Leu Val Asp Ala Leu Arg Tyr Gly Gln
            260                 265                 270
Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn
        275                 280                 285
His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile
    290                 295                 300
Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu
305                 310                 315                 320
Trp Leu Asn Lys Asp Lys Leu Lys Asn Leu Ile Trp Ile His Asp Asn
                325                 330                 335
Leu Ala Gly Gly Ser Thr Ser Ile Val Tyr Tyr Asp Ser Asp Glu Met
            340                 345                 350
Ile Phe Val Arg Thr Gly Tyr Gly Ser Lys Pro Gly Leu Ile Thr Tyr
        355                 360                 365
Ile Asn Leu Gly Ser Ser Lys Val Gly Arg Trp Val Tyr Val Pro Lys
    370                 375                 380
Phe Ala Gly Ala Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp
385                 390                 395                 400
Val Asp Lys Tyr Val Tyr Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro
                405                 410                 415
Ala Tyr Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr
            420                 425                 430
Cys Gly Val Gly
        435
```

<210> 33
<211> 1311
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetically generated oligonucleotide

<400> 33

```
atggccaagt actccgagct ggaagagggc ggggtcataa tgcaggcgtt ctactgggac      60
gtgccttcag gaggaatatg gtgggacaca atacggcaga agataccgga gtggtacgat     120
gccggaatct ccgcaatatg gattcctccc gcgagcaagg gtatgagcgg cggctattcg     180
atgggctacg acccctacga ttattttgac ctcggtgagt actaccagaa gggaacggtg     240
gaaacgaggt tcggctcaaa gcaggagctc ataaacatga taaacacggc ccatgcctac     300
ggcataaagg tcatagcgga catcgtcata aaccaccgcg caggcggaga cctcgagtgg     360
aacccgttcg ttggggacta cacctggacg gacttctcaa aggtggcctc gggcaaatat     420
actgccaact acctcgactt ccacccgaac gagctccatg cgggcgattc cggaacattt     480
ggaggctatc ccgacatatg ccacgacaag agctgggacc agtactggct ctgggccagc     540
caggagagct acgcggcata tctcaggagc atcggcatcg atgcctggcg ctttgactac     600
gtgaagggct acggagcgtg ggtcgtcaag gactggctca actggtgggg cggctgggcc     660
gttggcgagt actgggacac caacgttgat gcactcctca actgggccta ctcgagcggc     720
gccaaggtct tcgactttcc gctctactac aagatggacg cggcctttga caacaagaac     780
attcccgcac tcgtcgaggc cctcaagaac gggggcacag tcgtcagccg cgacccgttt     840
aaggccgtaa ccttcgttgc aaaccacgac accgatataa tctggaccaa gtaccttgct     900
tatgctttca tcctcaccta cgaaggccag cccgtcatat tctaccgcga ctacgaggag     960
tggctcaaca aggacaggtt gaacaacctc atatggatac acgaccacct cgcaggtgga    1020
agcaccgaca tagtctacta cgataacgat gaactcatct tcgtcaggaa cggctacggg    1080
gacaagccgg ggcttataac ctacatcaac ctaggctcga gcaaggccgg aaggtgggtt    1140
tacgttccga agttcgcagg ctcgtgcata cacgagtaca ccggcaatct cggcggctgg    1200
gtggacaagt gggtggactc aagcggctgg gtctacctcg aggctcctgc ccacgacccg    1260
gccaacggcc agtacggcta ctccgtctgg agctactgcg gtgttggctg a             1311
```

<210> 34
<211> 436
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetically generated peptide

<400> 34

```
Met Ala Lys Tyr Ser Glu Leu Glu Glu Gly Gly Val Ile Met Gln Ala
 1               5                  10                  15
```

```
Phe Tyr Trp Asp Val Pro Ser Gly Gly Ile Trp Trp Asp Thr Ile Arg
            20                  25                  30
Gln Lys Ile Pro Glu Trp Tyr Asp Ala Gly Ile Ser Ala Ile Trp Ile
        35                  40                  45
Pro Pro Ala Ser Lys Gly Met Ser Gly Gly Tyr Ser Met Gly Tyr Asp
    50                  55                  60
Pro Tyr Asp Tyr Phe Asp Leu Gly Glu Tyr Tyr Gln Lys Gly Thr Val
65                  70                  75                  80
Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Ile Asn Met Ile Asn Thr
                85                  90                  95
Ala His Ala Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His
            100                 105                 110
Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Gly Asp Tyr Thr
        115                 120                 125
Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr
    130                 135                 140
Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe
145                 150                 155                 160
Gly Gly Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp
                165                 170                 175
Leu Trp Ala Ser Gln Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly
            180                 185                 190
Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp Val
        195                 200                 205
Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr
    210                 215                 220
Trp Asp Thr Asn Val Asp Ala Leu Leu Asn Trp Ala Tyr Ser Ser Gly
225                 230                 235                 240
Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Ala Ala Phe
                245                 250                 255
Asp Asn Lys Asn Ile Pro Ala Leu Val Glu Ala Leu Lys Asn Gly Gly
            260                 265                 270
Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn
        275                 280                 285
His Asp Thr Asp Ile Ile Trp Thr Lys Tyr Leu Ala Tyr Ala Phe Ile
    290                 295                 300
Leu Thr Tyr Glu Gly Gln Pro Val Ile Phe Tyr Arg Asp Tyr Glu Glu
305                 310                 315                 320
Trp Leu Asn Lys Asp Arg Leu Asn Asn Leu Ile Trp Ile His Asp His
                325                 330                 335
Leu Ala Gly Gly Ser Thr Asp Ile Val Tyr Tyr Asp Asn Asp Glu Leu
            340                 345                 350
Ile Phe Val Arg Asn Gly Tyr Gly Asp Lys Pro Gly Leu Ile Thr Tyr
        355                 360                 365
Ile Asn Leu Gly Ser Ser Lys Ala Gly Arg Trp Val Tyr Val Pro Lys
    370                 375                 380
Phe Ala Gly Ser Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp
385                 390                 395                 400
Val Asp Lys Trp Val Asp Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro
                405                 410                 415
Ala His Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr
            420                 425                 430
Cys Gly Val Gly
        435
```

<210> 35
<211> 1311
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetically generated oligonucleotide

<400> 35
atggccaagt actccgagct ggaagagggc ggcgttataa tgcaggcctt ctactgggac 60
gtcccaggtg gaggaatctg gtgggacacc atcaggagca agataccgga gtggtacgag 120
gcgggaatat ccgccatttg gattcccccg gcgagcaagg gcatgggcgg cgcctattcg 180
atgggctacg acccctacga cttctttgac ctcggtgagt acgaccagaa gggaacggta 240

gagacgcgct ttggctccaa gcaggagctc gtgaacatga taaacaccgc ccacgcctac 300
ggcatcaagg tcatcgcaga catagtaatc aaccaccgcg ccggaggaga ccttgagtgg 360
aaccccttcg tcaatgacta cacctggacg gacttctcga aggtcgcttc cggcaagtac 420
acggccaact acctcgactt ccaccccaac gaggtcaagt gctgtgacga gggcacattt 480
ggaggcttcc cagacatagc ccacgagaag agctgggacc agcactggct ctgggcgagc 540
gatgagagct acgccgccta cctaaggagc atcggcgttg atgcctggcg cttcgactac 600
gtcaagggct atgctccctg ggtcgtcaag gactggctga actggtgggg aggctgggcg 660
gttggagagt actgggacac caacgtcgac gctgttctca actgggcata ctcgagcggt 720
gccaaggtct ttgacttcgc cctctactac aagatggacg cggcctttga caacaagaac 780
attcccgcac tcgtcgaggc cctcaagaac gggggcacag tcgtcagccg cgacccgttt 840
aaggccgtaa ccttcgttgc aaaccacgac accgatataa tctggaacaa gtatccagcc 900
tacgcgttca tcctcaccta cgagggccag ccgacaatat tctaccgcga ctacgaggag 960
tggctcaaca aggataagct caagaacctc atctggatac atgacaacgt cgccggagga 1020
agcaccgaca tagtctacta cgataacgat gaactcatct tcgtcaggaa cggctacggg 1080
gacaagccgg ggcttataac ctacatcaac ctaggctcga gcaaggccgg aaggtgggtt 1140
tacgttccga agttcgcagg ctcgtgcata cacgagtaca ccggcaatct cggcggctgg 1200
gtggacaagt gggtggactc aagcggctgg gtctacctcg aggctcctgc ccacgacccg 1260
gccaacggcc agtacggcta ctccgtctgg agctactgcg gtgttgggtg a 1311

<210> 36
<211> 436
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetically generated peptide

<400> 36
Met Ala Lys Tyr Ser Glu Leu Glu Glu Gly Gly Val Ile Met Gln Ala
 1   5   10   15
Phe Tyr Trp Asp Val Pro Gly Gly Gly Ile Trp Trp Asp Thr Ile Arg
            20   25   30
Ser Lys Ile Pro Glu Trp Tyr Glu Ala Gly Ile Ser Ala Ile Trp Ile
        35   40   45
Pro Pro Ala Ser Lys Gly Met Gly Gly Ala Tyr Ser Met Gly Tyr Asp
    50   55   60
Pro Tyr Asp Phe Phe Asp Leu Gly Glu Tyr Asp Gln Lys Gly Thr Val
65   70   75   80
Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Val Asn Met Ile Asn Thr
                85   90   95
Ala His Ala Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His
            100   105   110
Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Asn Asp Tyr Thr
        115   120   125
Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr
    130   135   140
Leu Asp Phe His Pro Asn Glu Val Lys Cys Cys Asp Glu Gly Thr Phe
145   150   155   160
Gly Gly Phe Pro Asp Ile Ala His Glu Lys Ser Trp Asp Gln His Trp
                165   170   175
Leu Trp Ala Ser Asp Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly
            180   185   190
Val Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Ala Pro Trp Val
        195   200   205
Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr
    210   215   220
Trp Asp Thr Asn Val Asp Ala Val Leu Asn Trp Ala Tyr Ser Ser Gly
225   230   235   240
Ala Lys Val Phe Asp Phe Ala Leu Tyr Tyr Lys Met Asp Ala Ala Phe
                245   250   255
Asp Asn Lys Asn Ile Pro Ala Leu Val Glu Ala Leu Lys Asn Gly Gly
            260   265   270
Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn
        275   280   285
His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile
    290   295   300
Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu
305   310   315   320

Trp Leu Asn Lys Asp Lys Leu Lys Asn Leu Ile Trp Ile His Asp Asn
325 330 335
Val Ala Gly Gly Ser Thr Asp Ile Val Tyr Tyr Asp Asn Asp Glu Leu
340 345 350
Ile Phe Val Arg Asn Gly Tyr Gly Asp Lys Pro Gly Leu Ile Thr Tyr
355 360 365
Ile Asn Leu Gly Ser Ser Lys Ala Gly Arg Trp Val Tyr Val Pro Lys
370 375 380
Phe Ala Gly Ser Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp
385 390 395 400
Val Asp Lys Trp Val Asp Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro
405 410 415
Ala His Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr
420 425 430
Cys Gly Val Gly
435

<210> 37
<211> 1311
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetically generated oligonucleotide

<400> 37
atggccaagt acctggagct cgaagagggc ggggtcataa tgcaggcgtt ctactgggac 60
gtgccttcag gaggaatatg gtgggacaca atacggcaga agataccgga gtggtacgat 120
gccggaatct ccgcaatatg gattcccccg gcgagcaagg gcatgggcgg cgcctattcg 180
atgggctacg acccctacga cttctttgac ctcggtgagt acgaccagaa gggaacggta 240
gagacgcgct ttggctccaa gcaggagctc gtgaacatga taaacaccgc ccacgcctat 300
ggcatgaagg taatagccga tatagtcatc aaccaccgcg ccggcggtga cctggagtgg 360
aacccccttcg tgaacgacta tacctggacc gacttctcaa aggtcgcgtc gggtaaatac 420
acggccaact acctcgactt ccacccgaac gagctccatg cgggcgattc cggaacattt 480
ggaggctatc ccgacatatg ccacgacaag agctgggacc agtactggct ctgggccagc 540
caggagagct acgcggcata tctcaggagc atcggcatcg atgcctggcg ctttgactac 600
gtgaagggct acggagcgcg ggtcgtcaag gactggctca actggtgggg cggctgggcc 660
gttggcgagt actgggacac caacgttgat gcactcctca actgggccta ctcgagcggc 720
gccaaggtct tcgacttccc gctctactac aagatggatg aggcctttga caacaaaaac 780
attccagcgc tcgtctctgc ccttcagaac ggccagactg ttgtctcccg cgacccgttc 840
aaggccgtaa cctttgtagc aaaccacgac accgatataa tctggaacaa gtatccagcc 900
tacgcgttca tcctcaccta cgagggccag ccgacaatat tctatcgcga ctacgaggag 960
tggctcaaca aggataagct caagaacctc atctggatac atgacaacct cgccggagga 1020
agcactgaca tcgtttacta cgacaacgac gagctgatat tcgtgagaaa cggctacgga 1080
agcaagccgg gactgataac atacatcaac ctcgcctcaa gcaaagccgg aaggtgggtt 1140
tacgttccga agttcgcagg ctcgtgcata cacgagtaca ccggcaatct cggcggctgg 1200
gtggacaagt gggtggactc aagcggctgg gtctacctcg aggctcctgc ccacgacccg 1260
gccaacggcc agtacggcta ctccgtctgg agctactgcg gggtggggtg a 1311

<210> 38
<211> 436
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetically generated peptide

<400> 38
Met Ala Lys Tyr Leu Glu Leu Glu Glu Gly Gly Val Ile Met Gln Ala
1 5 10 15
Phe Tyr Trp Asp Val Pro Ser Gly Gly Ile Trp Trp Asp Thr Ile Arg
20 25 30
Gln Lys Ile Pro Glu Trp Tyr Asp Ala Gly Ile Ser Ala Ile Trp Ile
35 40 45
Pro Pro Ala Ser Lys Gly Met Gly Gly Ala Tyr Ser Met Gly Tyr Asp
50 55 60
Pro Tyr Asp Phe Phe Asp Leu Gly Glu Tyr Asp Gln Lys Gly Thr Val
65 70 75 80

```
Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Val Asn Met Ile Asn Thr
                85                  90                  95
Ala His Ala Tyr Gly Met Lys Val Ile Ala Asp Ile Val Ile Asn His
            100                 105                 110
Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Asn Asp Tyr Thr
        115                 120                 125
Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr
    130                 135                 140
Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe
145                 150                 155                 160
Gly Gly Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp
                165                 170                 175
Leu Trp Ala Ser Gln Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly
            180                 185                 190
Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala Arg Val
        195                 200                 205
Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr
    210                 215                 220
Trp Asp Thr Asn Val Asp Ala Leu Leu Asn Trp Ala Tyr Ser Ser Gly
225                 230                 235                 240
Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Glu Ala Phe
                245                 250                 255
Asp Asn Lys Asn Ile Pro Ala Leu Val Ser Ala Leu Gln Asn Gly Gln
            260                 265                 270
Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn
        275                 280                 285
His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile
    290                 295                 300
Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu
305                 310                 315                 320
Trp Leu Asn Lys Asp Lys Leu Lys Asn Leu Ile Trp Ile His Asp Asn
                325                 330                 335
Leu Ala Gly Gly Ser Thr Asp Ile Val Tyr Tyr Asp Asn Asp Glu Leu
            340                 345                 350
Ile Phe Val Arg Asn Gly Tyr Gly Ser Lys Pro Gly Leu Ile Thr Tyr
        355                 360                 365
Ile Asn Leu Ala Ser Ser Lys Ala Gly Arg Trp Val Tyr Val Pro Lys
    370                 375                 380
Phe Ala Gly Ser Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp
385                 390                 395                 400
Val Asp Lys Trp Val Asp Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro
                405                 410                 415
Ala His Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr
            420                 425                 430
Cys Gly Val Gly
        435
```

<210> 39
<211> 1311
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetically generated oligonucleotide

<400> 39

```
atggccaagt acctggagct cgaagagggc ggggtcataa tgcaggcgtt ctactgggac      60
gtgccttcag gaggaatatg gtgggacaca atacggcaga agataccgga gtggtacgat     120
gccggaatct ccgcaatatg gattcctccc gcgagcaggg gtatgagcgg cggctattcg     180
atgggctacg acccctacga ttattttgac ctcggtgagt actaccagaa gggaacggtg     240
gaaacgaggt tcggctcaaa gcaggagctc ataaacatga taaacaccgc ccacgcctat     300
ggcatgaagg taatagccga tatagtcatc aaccaccgcg ccggcggtga cctggagtgg     360
aacccttcg tgaacgacta tacctggacc gacttctcaa aggtcgcgtc gggtaaatac     420
acggccaact acctcgactt ccacccgaac gagctccatg cgggcgattc cggaacattt     480
ggaggctatc ccgacatatg ccacgacaag agctgggacc agtactggct ctgggccagc     540
caggagagct acgcggcata tctcaggagc atcggtatcg atgcctggcg ctttgactac     600
gtgaagggct acggagcgtg ggtcgtcaag gactggctca actggtgggg cggctgggcc     660
gttggcgagt actgggaccc caacgttgat gccctcctcc cctgggccta ctcgagcggc     720
```

gccaaggtct tcgacttccc gctctactac aagatggatg aggcctttga caacaaaaac 780
attccagcgc tcgtctctgc ccttcagaac ggccagactg ttgtctcccg cgacccgttc 840
aaggccgtaa cctttgtagc caaccacgat accgatataa tctggaacaa gtatccagcc 900
tacgcgttca tcctcaccta cgagggccag ccgacaatat tctaccgcga ctacgaggag 960
tggctcaaca aggataagct caagaacctc atctggatac atgacaacct cgccggagga 1020
agcaccgaca tagtctacta cgataacgat gaactcatct tcgtcaggaa cggctacggg 1080
gacaagccgg ggcttataac ctacatcaac ctaggctcga gcaaggccgg aaggtgggtc 1140
tacgttccga agttcgcggg agcgtgcatc cacgagtaca ccggcaacct cggcggctgg 1200
gtggacaagt gggtggactc aagcgggtgg gtgtacctcg aggcccctgc ccacgacccg 1260
gccaacggct attacggcta ctccgtctgg agctactgcg gggtgggctg a 1311

<210> 40
<211> 436
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetically generated peptide

<400> 40
Met Ala Lys Tyr Leu Glu Leu Glu Glu Gly Gly Val Ile Met Gln Ala
 1   5   10   15
Phe Tyr Trp Asp Val Pro Ser Gly Gly Ile Trp Trp Asp Thr Ile Arg
            20   25   30
Gln Lys Ile Pro Glu Trp Tyr Asp Ala Gly Ile Ser Ala Ile Trp Ile
        35   40   45
Pro Pro Ala Ser Arg Gly Met Ser Gly Gly Tyr Ser Met Gly Tyr Asp
    50   55   60
Pro Tyr Asp Tyr Phe Asp Leu Gly Glu Tyr Tyr Gln Lys Gly Thr Val
65   70   75   80
Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Ile Asn Met Ile Asn Thr
                85   90   95
Ala His Ala Tyr Gly Met Lys Val Ile Ala Asp Ile Val Ile Asn His
            100   105   110
Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Asn Asp Tyr Thr
        115   120   125
Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr
    130   135   140
Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe
145   150   155   160
Gly Gly Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp
                165   170   175
Leu Trp Ala Ser Gln Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly
            180   185   190
Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp Val
        195   200   205
Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr
    210   215   220
Trp Asp Pro Asn Val Asp Ala Leu Leu Pro Trp Ala Tyr Ser Ser Gly
225   230   235   240
Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Glu Ala Phe
                245   250   255
Asp Asn Lys Asn Ile Pro Ala Leu Val Ser Ala Leu Gln Asn Gly Gln
            260   265   270
Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn
        275   280   285
His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile
    290   295   300
Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu
305   310   315   320
Trp Leu Asn Lys Asp Lys Leu Lys Asn Leu Ile Trp Ile His Asp Asn
                325   330   335
Leu Ala Gly Gly Ser Thr Asp Ile Val Tyr Tyr Asp Asn Asp Glu Leu
            340   345   350
Ile Phe Val Arg Asn Gly Tyr Gly Asp Lys Pro Gly Leu Ile Thr Tyr
        355   360   365
Ile Asn Leu Gly Ser Ser Lys Ala Gly Arg Trp Val Tyr Val Pro Lys
    370   375   380

Phe Ala Gly Ala Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp
385 390 395 400
Val Asp Lys Trp Val Asp Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro
405 410 415
Ala His Asp Pro Ala Asn Gly Tyr Tyr Gly Tyr Ser Val Trp Ser Tyr
420 425 430
Cys Gly Val Gly
435

<210> 41
<211> 1311
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetically generated oligonucleotide

<400> 41
atggccaagt acctggagct cgaagagggc ggggtcataa tgcaggcgtt ctactgggac 60
gtgccttcag gaggaatatg gtgggacaca atacggcaga agataccgga gtggtacgat 120
gccggaatct ccgcaatatg gattcctccc gcgagcaagg gtatgagcgg cggctattcg 180
atgggctacg acccctacga ttattttgac ctcggtgagt actaccagaa gggaacggtg 240
gaaacgaggt tcggctcaaa gcaggagctc ataaacatga taaacacggc ccatgcctac 300
ggcataaagg tcatagcgga catcgtcata aaccaccgcg caggcggaga cctcgagtgg 360
aacccgttcg ttggggacta cacctggacg gacttctcaa aggtggcctc gggcaaatat 420
actgccaact acctcgactt ccacccgaac gagctccatg cgggcgattc cggaacattt 480
ggaggctatc ccgacatatg ccacgacaag agctgggacc agtactggct ctgggccagc 540
caggagagct acgcggcata tctcaggagc atcggcatcg atgcctggcg ctttgactac 600
gtgaagggct acggagcgtg ggtcgtcaag gactggctca actggtgggg cggctgggcc 660
gttggcgagt actgggacac caacgttgat gcactcctca actgggccta ctcgagcggc 720
gccaaggtct tcgacttccc gctctactac aagatggacg cggcctttga caacaagaac 780
attcccgcac tcgtcgaggc cctcaagaac gggggcacag tcgtcagccg cgacccgttt 840
aaggccgtaa ccttcgttgc aaaccacgac accgatataa tctggaacaa gtatccagcc 900
tacgcgttca tcctcaccta cgagggccag ccgacaatat tctaccgcga ctacgaggag 960
tggctcaaca aggataagct caagaacctc atctggatac atgacaacct cgccggagga 1020
agcacgagca tagtttacta cgacagcgac gagatgatct tcgtgaggaa cggctatgga 1080
agcaagcctg gccttataac ttacatcaac ctcggctcga gcaaggttgg aaggtgggtt 1140
tatgtgccga agttcgcggg cgcgtgcatc cacgagtata ctggtaacct cggaggctgg 1200
gtagacaagt acgtctactc aagcggctgg gtctatctcg aagctccagc ttacgaccct 1260
gccaacgggc agtatggcta ctccgtgtgg agctactgcg gtgttgggtg a 1311

<210> 42
<211> 436
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetically generated peptide

<400> 42
Met Ala Lys Tyr Leu Glu Leu Glu Glu Gly Gly Val Ile Met Gln Ala
1 5 10 15
Phe Tyr Trp Asp Val Pro Ser Gly Gly Ile Trp Trp Asp Thr Ile Arg
20 25 30
Gln Lys Ile Pro Glu Trp Tyr Asp Ala Gly Ile Ser Ala Ile Trp Ile
35 40 45
Pro Pro Ala Ser Lys Gly Met Ser Gly Gly Tyr Ser Met Gly Tyr Asp
50 55 60
Pro Tyr Asp Tyr Phe Asp Leu Gly Glu Tyr Tyr Gln Lys Gly Thr Val
65 70 75 80
Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Ile Asn Met Ile Asn Thr
85 90 95
Ala His Ala Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His
100 105 110
Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Gly Asp Tyr Thr
115 120 125
Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr
130 135 140

```
Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe
145                 150                 155                 160
Gly Gly Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp
                165                 170                 175
Leu Trp Ala Ser Gln Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly
            180                 185                 190
Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp Val
        195                 200                 205
Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr
    210                 215                 220
Trp Asp Thr Asn Val Asp Ala Leu Leu Asn Trp Ala Tyr Ser Ser Gly
225                 230                 235                 240
Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Ala Ala Phe
                245                 250                 255
Asp Asn Lys Asn Ile Pro Ala Leu Val Glu Ala Leu Lys Asn Gly Gly
            260                 265                 270
Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn
        275                 280                 285
His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile
    290                 295                 300
Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu
305                 310                 315                 320
Trp Leu Asn Lys Asp Lys Leu Lys Asn Leu Ile Trp Ile His Asp Asn
                325                 330                 335
Leu Ala Gly Gly Ser Thr Ser Ile Val Tyr Tyr Asp Ser Asp Glu Met
            340                 345                 350
Ile Phe Val Arg Asn Gly Tyr Gly Ser Lys Pro Gly Leu Ile Thr Tyr
        355                 360                 365
Ile Asn Leu Gly Ser Ser Lys Val Gly Arg Trp Val Tyr Val Pro Lys
    370                 375                 380
Phe Ala Gly Ala Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp
385                 390                 395                 400
Val Asp Lys Tyr Val Tyr Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro
                405                 410                 415
Ala Tyr Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr
            420                 425                 430
Cys Gly Val Gly
        435
```

<210> 43
<211> 1311
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetically generated oligonucleotide

<400> 43

```
atggccaagt actccgagct ggaagagggc ggcgttataa tgcaggcctt ctactgggac      60
gtcccaggtg gaggaatctg gtgggacacc atcaggagca agataccgga gtggtacgag     120
gcgggaatat ccgccatttg gattcccccg gcgagcaagg gcatgggcgg cgcctattcg     180
atgggctacg acccctacga cttctttgac ctcggtgagt acgaccagaa gggaacggta     240
gagacgcgct ttggctccaa gcaggagctc gtgaacatga taaacacggc ccatgcctac     300
ggcataaagg tcatagcgga catcgtcata aaccaccgcg caggcggaga cctcgagtgg     360
aacccgttcg ttggggacta cacctggacg gacttctcaa aggtggcctc gggcaaatat     420
actgccaact acctcgactt ccaccccaac gaggtcaagt gctgtgacga gggcacattt     480
ggaggcttcc cagacatagc ccacgagaag agctgggacc agcactggct ctgggcgagc     540
gatgagagct acgccgccta cctaaggagc atcggcgttg atgcctggcg cttcgactac     600
gtcaagggct acggagcgtg ggtcgtcaag gactggctgg actggtgggg aggctgggcc     660
gtcggggagt actgggacac aaacgttgat gcactgctca actgggccta ctcgagcgat     720
gcaaaagtct tcgacttccc gctctactac aagatggatg aggcctttga caacaaaaac     780
attccagcgc tcgtctctgc ccttcagaac ggccagactg ttgtctcccg cgacccgttc     840
aaggccgtaa cctttgtagc aaaccacgac accgatataa tctggaacaa gtatccagcc     900
tacgcgttca tcctcaccta cgagggccag ccgacaatat tctaccgcga ctacgaggag     960
tggctcaaca aggataagct caagaacctc atctggatac atgacaacct cgtcggagga    1020
agcacgagca tagtttacta cgacagcgac gagatgatct tcgtgaggaa cggctatgga    1080
agcaagcctg gccttataac ttacatcaac ctcggctcga gcaaggttgg aaggtgggtt    1140
tacgttccga agttcgcagg ctcgtgcata cacgagtaca ccggcaatct cggcggctgg    1200
```

gtggacaagt gggtggactc aagcggctgg gtctacctcg aggctcctgc ccacgacccg 1260
gccaacggcc agtacggcta ctccgtctgg agctactgcg gtgttggctg a 1311

<210> 44
<211> 436
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetically generated peptide

<400> 44
Met Ala Lys Tyr Ser Glu Leu Glu Glu Gly Gly Val Ile Met Gln Ala
1 5 10 15
Phe Tyr Trp Asp Val Pro Gly Gly Gly Ile Trp Trp Asp Thr Ile Arg
20 25 30
Ser Lys Ile Pro Glu Trp Tyr Glu Ala Gly Ile Ser Ala Ile Trp Ile
35 40 45
Pro Pro Ala Ser Lys Gly Met Gly Gly Ala Tyr Ser Met Gly Tyr Asp
50 55 60
Pro Tyr Asp Phe Phe Asp Leu Gly Glu Tyr Asp Gln Lys Gly Thr Val
65 70 75 80
Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Val Asn Met Ile Asn Thr
85 90 95
Ala His Ala Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His
100 105 110
Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Gly Asp Tyr Thr
115 120 125
Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr
130 135 140
Leu Asp Phe His Pro Asn Glu Val Lys Cys Cys Asp Glu Gly Thr Phe
145 150 155 160
Gly Gly Phe Pro Asp Ile Ala His Glu Lys Ser Trp Asp Gln His Trp
165 170 175
Leu Trp Ala Ser Asp Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly
180 185 190
Val Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp Val
195 200 205
Val Lys Asp Trp Leu Asp Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr
210 215 220
Trp Asp Thr Asn Val Asp Ala Leu Leu Asn Trp Ala Tyr Ser Ser Asp
225 230 235 240
Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Glu Ala Phe
245 250 255
Asp Asn Lys Asn Ile Pro Ala Leu Val Ser Ala Leu Gln Asn Gly Gln
260 265 270
Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn
275 280 285
His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile
290 295 300
Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu
305 310 315 320
Trp Leu Asn Lys Asp Lys Leu Lys Asn Leu Ile Trp Ile His Asp Asn
325 330 335
Leu Val Gly Gly Ser Thr Ser Ile Val Tyr Tyr Asp Ser Asp Glu Met
340 345 350
Ile Phe Val Arg Asn Gly Tyr Gly Ser Lys Pro Gly Leu Ile Thr Tyr
355 360 365
Ile Asn Leu Gly Ser Ser Lys Val Gly Arg Trp Val Tyr Val Pro Lys
370 375 380
Phe Ala Gly Ser Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp
385 390 395 400
Val Asp Lys Trp Val Asp Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro
405 410 415
Ala His Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr
420 425 430
Cys Gly Val Gly
435

<210> 45
<211> 1311
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetically generated oligonucleotide

<400> 45
atggccaagt actccgacct ggaagagggc ggcgttataa tgcaggcctt ctactgggac 60
gtcccaggtg gaggaatctg gtgggacacc atcaggagca agataccgga gtggtacgag 120
gcgggaatat ccgccatttg gattcccccg gcgagcaagg gcatgggcgg cgcctattcg 180
atgggctacg acccctacga cttctttgac ctcggtgagt acgaccagaa gggaacggta 240
gagacgcgct ttggctccaa gcaggagctc gtgaacatga taaacacggc ccatgcctac 300
ggcataaagg tcatagcgga catcgtcata aaccaccgcg caggcggaga cctcgagtgg 360
aacccgttcg ttggggacta cacctggacg gacttctcaa aggtggcctc gggcaaatat 420
actgccaact acctcgactt ccaccccaac gaggtcaagt gctgtgacga gggcacattt 480
ggaggcttcc cagacatagc ccacgagaag agctgggacc agcactggct ctgggcgagc 540
gatgagagct acgccgccta cctaaggagc atcggcgttg atgcctggcg ctttgactac 600
gtgaagggct acggagcgtg ggtcgtcaag gactggctca actggtgggg cggctgggcc 660
gttggcgagt actgggacac caacgttgat gcactcctca actgggccta ctcgagcggc 720
gccaaggtct tcgacttccc gctctactac aagatggatg aggcctttga caacaaaaac 780
attccagcgc tcgtctctgc ccttcagaac ggccagactg ttgtctcccg cgacccgttc 840
aaggccgtaa cctttgtagc aaaccacgac accgatataa tctggaacaa gtatccagcc 900
tacgcgttca tcctcaccta cgagggccag ccgacaatat tctaccgcga ctacgaggag 960
tggctcaaca aggataagct caagaacctc atctggatac atgacaacct cgccggagga 1020
agcaccgaca tagtctacta cgataacgat gaactcatct tcgtcaggaa cggctacggg 1080
gacaagccgg ggcttataac ctacatcaac ctaggctcga gcaaggccgg aaggtgggtt 1140
tatgtgccga agttcgcggg cgcgtgcatc cacgagtata ctggtaacct cggaggctgg 1200
gtagacaagt acgtctactc aagcggctgg gtctatctcg aagctccagc ttacgaccct 1260
gccaacgggc agtatggcta ctccgtgtgg agctattgcg gtgttgggtg a 1311

<210> 46
<211> 436
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetically generated peptide

<400> 46
Met Ala Lys Tyr Ser Asp Leu Glu Glu Gly Gly Val Ile Met Gln Ala
 1   5   10   15
Phe Tyr Trp Asp Val Pro Gly Gly Gly Ile Trp Trp Asp Thr Ile Arg
            20   25   30
Ser Lys Ile Pro Glu Trp Tyr Glu Ala Gly Ile Ser Ala Ile Trp Ile
        35   40   45
Pro Pro Ala Ser Lys Gly Met Gly Gly Ala Tyr Ser Met Gly Tyr Asp
    50   55   60
Pro Tyr Asp Phe Phe Asp Leu Gly Glu Tyr Asp Gln Lys Gly Thr Val
65   70   75   80
Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Val Asn Met Ile Asn Thr
                85   90   95
Ala His Ala Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His
            100   105   110
Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Gly Asp Tyr Thr
        115   120   125
Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr
    130   135   140
Leu Asp Phe His Pro Asn Glu Val Lys Cys Cys Asp Glu Gly Thr Phe
145   150   155   160
Gly Gly Phe Pro Asp Ile Ala His Glu Lys Ser Trp Asp Gln His Trp
                165   170   175
Leu Trp Ala Ser Asp Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly
            180   185   190
Val Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp Val
        195   200   205

```
Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr
    210                 215                 220
Trp Asp Thr Asn Val Asp Ala Leu Leu Asn Trp Ala Tyr Ser Ser Gly
225                 230                 235                 240
Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Glu Ala Phe
                245                 250                 255
Asp Asn Lys Asn Ile Pro Ala Leu Val Ser Ala Leu Gln Asn Gly Gln
            260                 265                 270
Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn
        275                 280                 285
His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile
    290                 295                 300
Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu
305                 310                 315                 320
Trp Leu Asn Lys Asp Lys Leu Lys Asn Leu Ile Trp Ile His Asp Asn
                325                 330                 335
Leu Ala Gly Gly Ser Thr Asp Ile Val Tyr Tyr Asp Asn Asp Glu Leu
            340                 345                 350
Ile Phe Val Arg Asn Gly Tyr Gly Asp Lys Pro Gly Leu Ile Thr Tyr
        355                 360                 365
Ile Asn Leu Gly Ser Ser Lys Ala Gly Arg Trp Val Tyr Val Pro Lys
    370                 375                 380
Phe Ala Gly Ala Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp
385                 390                 395                 400
Val Asp Lys Tyr Val Tyr Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro
                405                 410                 415
Ala Tyr Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr
            420                 425                 430
Cys Gly Val Gly
        435
```

<210> 47
<211> 1311
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetically generated oligonucleotide

<400> 47
```
atggccaagt acaccgagct ggaagagggc ggcgttataa tgcaggcctt ctactgggac      60
gtcccaggtg gaggaatctg gtgggacacc atcaggagca agataccgga gtggtacgag     120
gcgggaatat ccgccatttg gattcccccg gcgagcaagg gcatgggcgg cgcctattcg     180
atgggctacg acccctacga cttctttgac ctcggtgagt acgaccagaa gggaacggta     240
gagacgcgct ttggctccaa gcaggagctc gtgaacatga taaacaccgc ccacgcctat     300
ggcatgaagg taatagccga tatagtcatc aaccaccgcg ccggcggtga cctggagtgg     360
aacccccttcg tgaacgacta tacctggacc gacttctcaa aggtcgcgtc gggtaaatac     420
acggccaact acctcgactt ccaccccaac gaggtcaagt gctgtgacga gggcacattt     480
ggaggcttcc cagacatagc ccacgagaag agctgggacc agcactggct ctgggcgagc     540
gatgagagct acgccgccta cctaaggagc atcggcgttg atgcctggcg ctttgactac     600
gtgaagggct acggagcgtg ggtcgtcaag gactggctca actggtgggg cggttgggcc     660
gttggcgagt actgggacac caacgttgat gcactcctca actgggccta ctcgagcggc     720
gccaaggtct tcgacttccc gctctactac aagatggatg aggcctttga caacaaaaac     780
attccagcgc tcgtctctgc ccttcagaac ggccagactg ttgtctcccg cgacccgttc     840
aaggccgtaa cctttgtagc aaaccacgac accgatataa tctggaacaa gtaccttgct     900
tatgctttca tcctcaccta cgaaggccag cccgtcatat tctaccgcga ctacgaggag     960
tggctcaaca aggacaggtt gaacaacctc atatggatac acgaccacct cgcaggtgga    1020
agcacgagca tagtttacta cgacagcgac gagatgatct tcgtgaggaa cggctatgga    1080
agcaagcctg gccttataac ttacatcaac ctcggctcga gcaaggttgg aaggtgggtt    1140
tacgttccga agttcgcagg cccgtgcata cacgagtaca ccggcaatct cggcggctgg    1200
gtggacaagt gggtggactc aagcggctgg gtctacctcg aggctcctgc ccacgacccg    1260
gccaacggcc agtacggcta ctccgtctgg agctactgcg gtgttgggta g              1311
```

<210> 48
<211> 436
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetically generated peptide

<400> 48
Met Ala Lys Tyr Thr Glu Leu Glu Glu Gly Gly Val Ile Met Gln Ala
 1   5   10   15
Phe Tyr Trp Asp Val Pro Gly Gly Gly Ile Trp Trp Asp Thr Ile Arg
            20   25   30
Ser Lys Ile Pro Glu Trp Tyr Glu Ala Gly Ile Ser Ala Ile Trp Ile
        35   40   45
Pro Pro Ala Ser Lys Gly Met Gly Gly Ala Tyr Ser Met Gly Tyr Asp
    50   55   60
Pro Tyr Asp Phe Phe Asp Leu Gly Glu Tyr Asp Gln Lys Gly Thr Val
65   70   75   80
Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Val Asn Met Ile Asn Thr
                85   90   95
Ala His Ala Tyr Gly Met Lys Val Ile Ala Asp Ile Val Ile Asn His
            100   105   110
Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Asn Asp Tyr Thr
        115   120   125
Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr
    130   135   140
Leu Asp Phe His Pro Asn Glu Val Lys Cys Cys Asp Glu Gly Thr Phe
145   150   155   160
Gly Gly Phe Pro Asp Ile Ala His Glu Lys Ser Trp Asp Gln His Trp
                165   170   175
Leu Trp Ala Ser Asp Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly
            180   185   190
Val Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp Val
        195   200   205
Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr
    210   215   220
Trp Asp Thr Asn Val Asp Ala Leu Leu Asn Trp Ala Tyr Ser Ser Gly
225   230   235   240
Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Glu Ala Phe
                245   250   255
Asp Asn Lys Asn Ile Pro Ala Leu Val Ser Ala Leu Gln Asn Gly Gln
            260   265   270
Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn
        275   280   285
His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Leu Ala Tyr Ala Phe Ile
    290   295   300
Leu Thr Tyr Glu Gly Gln Pro Val Ile Phe Tyr Arg Asp Tyr Glu Glu
305   310   315   320
Trp Leu Asn Lys Asp Arg Leu Asn Asn Leu Ile Trp Ile His Asp His
                325   330   335
Leu Ala Gly Gly Ser Thr Ser Ile Val Tyr Tyr Asp Ser Asp Glu Met
            340   345   350
Ile Phe Val Arg Asn Gly Tyr Gly Ser Lys Pro Gly Leu Ile Thr Tyr
        355   360   365
Ile Asn Leu Gly Ser Ser Lys Val Gly Arg Trp Val Tyr Val Pro Lys
    370   375   380
Phe Ala Gly Pro Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp
385   390   395   400
Val Asp Lys Trp Val Asp Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro
                405   410   415
Ala His Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr
            420   425   430
Cys Gly Val Gly
        435

<210> 49
<211> 387
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetically generated oligonucleotide

<400> 49
gtggtttatg acgatgtccg ctatgacctt tatgccgtag gcatgggccg tgtttatcat 60
gttcacgagc tcctgcttgg agccaaagcg cgtctctacc gttcccttct ggtcgtactc 120
accgaggtca aagaagtcgt aggggtcgta gcccatcgaa taggcgccgc ccatgccctt 180
gctcgccggg ggaatccata tcgccgaaat cccggcgctt gcccagtcgg gtatcttctg 240
ggctatcgtg tcccaccaga ttcctcccat ggggacgtcc cagtagaagg cctgcattat 300
gagcccgccc tcttcgagcc cggaatactt tgccataagt tacctcctac tagtagatta 360
aaattctgtt tcctgtgtga aattgtt 387

<210> 50
<211> 129
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetically generated peptide

<400> 50
Val Val Tyr Asp Asp Val Arg Tyr Asp Leu Tyr Ala Val Gly Met Gly
1 5 10 15
Arg Val Tyr His Val His Glu Leu Leu Gly Ala Lys Ala Arg Leu
20 25 30
Tyr Arg Ser Leu Leu Val Val Leu Thr Glu Val Lys Glu Val Val Gly
35 40 45
Val Val Ala His Arg Ile Gly Ala Ala His Ala Leu Ala Arg Arg Gly
50 55 60
Asn Pro Tyr Arg Arg Asn Pro Gly Ala Cys Pro Val Gly Tyr Leu Leu
65 70 75 80
Gly Tyr Arg Val Pro Pro Asp Ser Ser His Gly Asp Val Pro Val Glu
85 90 95
Gly Leu His Tyr Glu Pro Ala Leu Phe Glu Pro Gly Ile Leu Cys His
100 105 110
Lys Leu Pro Pro Thr Ser Arg Leu Lys Phe Cys Phe Leu Cys Glu Ile
115 120 125
Val

<210> 51
<211> 1311
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetically generated oligonucleotide

<400> 51
atggccaagt acctggagct cgaagagggc ggggtcataa tgcaggcgtt ctactgggac 60
gtgccttcag gaggaatatg gtgggacaca atacggcaga agataccgga gtggtacgat 120
gccggaatct ccgcaatatg gattcccccg gcgagcaagg gcatgggcgg cgcctattcg 180
atgggctacg acccctacga cttctttgac ctcggtgagt acgaccagaa gggaacggta 240
gagacgcgct ttggctccaa gcaggagctc gtgaacatga taaacaccgc ccacgcctat 300
ggcatgaagg taatagccga tatagtcatc aaccaccgcg ccggcggtga cctggagtgg 360
aacccttcg tgaacgacta tacctggacc gacttctcaa aggtcgcgtc gggtaaatac 420
acggccaact acctcgactt ccaccccaac gaggtcaagt gctgtgacga gggcacattt 480
ggaggcttcc cagacatagc ccacgagaag agctgggacc agcactggct ctgggcgagc 540
gatgagagct acgccgccta cctaaggagc atcggcgttg atgcctggcg ctttgactac 600
gtgaagggct acggagcgtg ggtcgtcaag gactggctca actggtgggg cggctgggcc 660
gttggcgagt actgggacac caacgttgat gcactcctca actgggccta ctcgagcggc 720
gccaaggtct tcgacttccc gctctactac aagatggatg aggcctttga caacaaaaac 780
attccagcgc tcgtctctgc ccttcagaac ggccagactg ttgtctcccg cgacccgttc 840
aaggccgtaa cctttgtagc aaaccacgac accgatataa tctggaacaa gtatccagcc 900
tacgcgttca tcctcaccta cgagggccag ccgacaatat tctaccgcga ctacgaggag 960
tggctcaaca aggataagct caagaacctc atctggatac atgacaacct cgccggagga 1020
agcactgaca tcgtttacta cgacaacgac gagctgatat tcgtgagaaa cggctacgga 1080
agcaagccgg gactgataac atacatcaac ctcgcctcaa gcaaagccgg aaggtgggtt 1140
tacgttccga agttcgcagg ctcgtgcata cacgagtaca ccggcaatct cggcggctgg 1200
gtggacaagt gggtggactc aagcggctgg gtctacctcg aggctcctgc ccacgacccg 1260

gccaacggcc agtacggcta ctccgtctgg agctattgcg gtgttggctg a 1311

<210> 52
<211> 436
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetically generated peptide

<400> 52

```
Met Ala Lys Tyr Leu Glu Leu Glu Glu Gly Gly Val Ile Met Gln Ala
 1               5                  10                  15
Phe Tyr Trp Asp Val Pro Ser Gly Gly Ile Trp Trp Asp Thr Ile Arg
            20                  25                  30
Gln Lys Ile Pro Glu Trp Tyr Asp Ala Gly Ile Ser Ala Ile Trp Ile
        35                  40                  45
Pro Pro Ala Ser Lys Gly Met Gly Gly Ala Tyr Ser Met Gly Tyr Asp
    50                  55                  60
Pro Tyr Asp Phe Phe Asp Leu Gly Glu Tyr Asp Gln Lys Gly Thr Val
65                  70                  75                  80
Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Val Asn Met Ile Asn Thr
                85                  90                  95
Ala His Ala Tyr Gly Met Lys Val Ile Ala Asp Ile Val Ile Asn His
            100                 105                 110
Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Asn Asp Tyr Thr
        115                 120                 125
Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr
    130                 135                 140
Leu Asp Phe His Pro Asn Glu Val Lys Cys Cys Asp Glu Gly Thr Phe
145                 150                 155                 160
Gly Gly Phe Pro Asp Ile Ala His Glu Lys Ser Trp Asp Gln His Trp
                165                 170                 175
Leu Trp Ala Ser Asp Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly
            180                 185                 190
Val Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp Val
        195                 200                 205
Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr
    210                 215                 220
Trp Asp Thr Asn Val Asp Ala Leu Leu Asn Trp Ala Tyr Ser Ser Gly
225                 230                 235                 240
Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Glu Ala Phe
                245                 250                 255
Asp Asn Lys Asn Ile Pro Ala Leu Val Ser Ala Leu Gln Asn Gly Gln
            260                 265                 270
Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn
        275                 280                 285
His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile
    290                 295                 300
Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu
305                 310                 315                 320
Trp Leu Asn Lys Asp Lys Leu Lys Asn Leu Ile Trp Ile His Asp Asn
                325                 330                 335
Leu Ala Gly Gly Ser Thr Asp Ile Val Tyr Tyr Asp Asn Asp Glu Leu
            340                 345                 350
Ile Phe Val Arg Asn Gly Tyr Gly Ser Lys Pro Gly Leu Ile Thr Tyr
        355                 360                 365
Ile Asn Leu Ala Ser Ser Lys Ala Gly Arg Trp Val Tyr Val Pro Lys
    370                 375                 380
Phe Ala Gly Ser Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp
385                 390                 395                 400
Val Asp Lys Trp Val Asp Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro
                405                 410                 415
Ala His Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr
            420                 425                 430
Cys Gly Val Gly
        435
```

<210> 53
<211> 1311
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetically generated oligonucleotide

<400> 53
atggccaagt actccgagct ggaagagggc ggcgttataa tgcaggcctt ctactgggac 60
gtcccaggtg gaggaatctg gtgggacacc atcaggagca agataccgga gtggtacgag 120
gcgggaatat ccgccatttg gattcccccg gcgagcaagg gcatgggcgg cgcctattcg 180
atgggctacg acccctacga cttctttgac ctcggtgagt acgaccagaa gggaacggta 240
gagacgcgct ttggctccaa gcaggagctc gtgaacatga taaacacggc ccatgcctac 300
ggcataaagg tcatagcgga catcgtcata aaccaccgca caggcggaga cctcgagtgg 360
aacccgttcg ttggggacta cacctggacg gacttctcaa aggtggcctc gggcaaatat 420
actgccaact acctcgactt ccaccccaac gaggtcaagt gctgtgacga gggcacattt 480
ggaggcttcc cagacatagc ccacgagaag agctgggacc agcactggct ctgggcgagc 540
gatgagagct acgccgccta cctaaggagc atcggcgttg atgcctggcg cttcgactac 600
gtcaagggct acggagcgtg ggtcgtcaag gactggctgg actggtgggg aggctgggcc 660
gtcggggagt actgggacac aaacgttgat gcactgctca actgggccta ctcgagcgat 720
gcaaaagtct tcgacttccc gctctactac aagatggatg aggcctttga caacaaaaac 780
attccagcgc tcgtctctgc ccttcagaac ggccagactg ttgtctcccg cgacccgttc 840
aaggccgtaa cctttgtagc aaaccacgac accgatataa tctggaacaa gtatccagcc 900
tacgcgttca tcctcaccta cgagggccag ccgacaatat tctaccgcga ctacgaggag 960
tggctcaaca aggataagct caagaacctc atctggatac atgacaacct cgccggagga 1020
agcactgaca tcgtttacta cgacaacgac gagctgatat tcgtgagaaa cggctacgga 1080
agcaagccgg gactgataac atacatcaac ctcgcctcaa gcaaagccgg aaggtgggtc 1140
tacgttccga agttcgcggg agcgtgcatc cacgagtaca ccggcaacct cggcggctgg 1200
gtggacaagt gggtggactc aagcgggtgg gtgtacctcg aggcccctgc ccacgacccg 1260
gccaacggct attacggcta ctccgtctgg agctactgcg gtgttggctg a 1311

<210> 54
<211> 436
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetically generated peptide

<400> 54
Met Ala Lys Tyr Ser Glu Leu Glu Glu Gly Gly Val Ile Met Gln Ala
 1   5   10   15
Phe Tyr Trp Asp Val Pro Gly Gly Gly Ile Trp Trp Asp Thr Ile Arg
            20   25   30
Ser Lys Ile Pro Glu Trp Tyr Glu Ala Gly Ile Ser Ala Ile Trp Ile
        35   40   45
Pro Pro Ala Ser Lys Gly Met Gly Gly Ala Tyr Ser Met Gly Tyr Asp
    50   55   60
Pro Tyr Asp Phe Phe Asp Leu Gly Glu Tyr Asp Gln Lys Gly Thr Val
65   70   75   80
Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Val Asn Met Ile Asn Thr
                85   90   95
Ala His Ala Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His
            100   105   110
Arg Thr Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Gly Asp Tyr Thr
        115   120   125
Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr
    130   135   140
Leu Asp Phe His Pro Asn Glu Val Lys Cys Cys Asp Glu Gly Thr Phe
145   150   155   160
Gly Gly Phe Pro Asp Ile Ala His Glu Lys Ser Trp Asp Gln His Trp
                165   170   175
Leu Trp Ala Ser Asp Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly
            180   185   190
Val Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp Val
        195   200   205
Val Lys Asp Trp Leu Asp Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr

210 215 220
Trp Asp Thr Asn Val Asp Ala Leu Leu Asn Trp Ala Tyr Ser Ser Asp
225 230 235 240
Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Glu Ala Phe
245 250 255
Asp Asn Lys Asn Ile Pro Ala Leu Val Ser Ala Leu Gln Asn Gly Gln
260 265 270
Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn
275 280 285
His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile
290 295 300
Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu
305 310 315 320
Trp Leu Asn Lys Asp Lys Leu Lys Asn Leu Ile Trp Ile His Asp Asn
325 330 335
Leu Ala Gly Gly Ser Thr Asp Ile Val Tyr Tyr Asp Asn Asp Glu Leu
340 345 350
Ile Phe Val Arg Asn Gly Tyr Gly Ser Lys Pro Gly Leu Ile Thr Tyr
355 360 365
Ile Asn Leu Ala Ser Ser Lys Ala Gly Arg Trp Val Tyr Val Pro Lys
370 375 380
Phe Ala Gly Ala Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp
385 390 395 400
Val Asp Lys Trp Val Asp Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro
405 410 415
Ala His Asp Pro Ala Asn Gly Tyr Tyr Gly Tyr Ser Val Trp Ser Tyr
420 425 430
Cys Gly Val Gly
435

<210> 55
<211> 1311
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetically generated oligonucleotide

<400> 55
atggccaagt acctggagct cgaggagggc ggggtcataa tgcaggcgtt ctactgggac 60
gtgccttcag gaggaatatg gtgggacaca atacggcaga agataccgga gtggtacgat 120
gccggaatct ccgcaatatg gattcccccg gcgagcaagg gcatgggcgg cgcctattcg 180
atgggctacg acccctacga cttctttgac ctcggtgagt acgaccagaa gggaacggta 240
gagacgcgct ttggctccaa gcaggagctc gtgaacatga taaacaccgc ccacgcctat 300
ggcatgaagg taatagccga tatagtcatc aaccaccgcg ccggcggtga cctggagtgg 360
aacccttcg tgaacgacta tacctggacc gacttctcaa aggtcgcgtc gggtaaatac 420
acggccaact acctcgactt ccacccgaac gagctccatg cgggcgattc cggaacattt 480
ggaggctatc ccgacatatg ccacgacaag agctgggacc agtactggct ctgggccagc 540
caggagagct acgcggcata tctcaggagc atcggcatcg atgcctggcg ctttgactac 600
gtgaagggct acggagcgtg ggtcgtcaag gactggctca actggtgggg cggctgggcc 660
gttggcgagt actgggacac caacgttgat gcactcctca actgggccta ctcgagcggc 720
gccaaggtct tcgacttccc gctctactac aagatggatg aggcctttga caacaaaaac 780
attccagcgc tcgtctctgc ccttcagaac ggccagactg ttgtctcccg cgacccgttc 840
aaggccgtaa cctttgtagc aaaccacgac accgatataa tctggaacaa gtaccttgct 900
tatgctttca tcctcaccta cgaaggccag cccgtcatat tctaccgcga ctacgaggag 960
tggctcaaca aggacaggtt gaacaacctc atatggatac acgaccacct cgcaggtgga 1020
agcacgagca tagtttacta cgacagcgac gagatgatct tcgtgaggaa cggctatgga 1080
agcaagcctg gccttataac ttacatcaac ctcggctcga gcaaggttgg aaggtgggtt 1140
tacgttccga agttcgcagg ctcgtgcata cacgagtaca ccggcaatct cggcggctgg 1200
gtggacaagt gggtggactc aagcggctgg gtctacctcg aggctcctgc ccacgacccg 1260
gccaacggcc agtacggcta ctccgtctgg agctattgcg gtgttggctg a 1311

<210> 56
<211> 436
<212> PRT
<213> Artificial Sequence

<220>

<223> Synthetically generated peptide

<400> 56

```
Met Ala Lys Tyr Leu Glu Leu Glu Glu Gly Gly Val Ile Met Gln Ala
 1               5                  10                  15
Phe Tyr Trp Asp Val Pro Ser Gly Gly Ile Trp Trp Asp Thr Ile Arg
            20                  25                  30
Gln Lys Ile Pro Glu Trp Tyr Asp Ala Gly Ile Ser Ala Ile Trp Ile
        35                  40                  45
Pro Pro Ala Ser Lys Gly Met Gly Gly Ala Tyr Ser Met Gly Tyr Asp
    50                  55                  60
Pro Tyr Asp Phe Phe Asp Leu Gly Glu Tyr Asp Gln Lys Gly Thr Val
65                  70                  75                  80
Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Val Asn Met Ile Asn Thr
                85                  90                  95
Ala His Ala Tyr Gly Met Lys Val Ile Ala Asp Ile Val Ile Asn His
            100                 105                 110
Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Asn Asp Tyr Thr
        115                 120                 125
Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr
    130                 135                 140
Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe
145                 150                 155                 160
Gly Gly Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp
                165                 170                 175
Leu Trp Ala Ser Gln Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly
            180                 185                 190
Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp Val
        195                 200                 205
Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr
    210                 215                 220
Trp Asp Thr Asn Val Asp Ala Leu Leu Asn Trp Ala Tyr Ser Ser Gly
225                 230                 235                 240
Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Glu Ala Phe
                245                 250                 255
Asp Asn Lys Asn Ile Pro Ala Leu Val Ser Ala Leu Gln Asn Gly Gln
            260                 265                 270
Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn
        275                 280                 285
His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Leu Ala Tyr Ala Phe Ile
    290                 295                 300
Leu Thr Tyr Glu Gly Gln Pro Val Ile Phe Tyr Arg Asp Tyr Glu Glu
305                 310                 315                 320
Trp Leu Asn Lys Asp Arg Leu Asn Asn Leu Ile Trp Ile His Asp His
                325                 330                 335
Leu Ala Gly Gly Ser Thr Ser Ile Val Tyr Tyr Asp Ser Asp Glu Met
            340                 345                 350
Ile Phe Val Arg Asn Gly Tyr Gly Ser Lys Pro Gly Leu Ile Thr Tyr
        355                 360                 365
Ile Asn Leu Gly Ser Ser Lys Val Gly Arg Trp Val Tyr Val Pro Lys
    370                 375                 380
Phe Ala Gly Ser Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp
385                 390                 395                 400
Val Asp Lys Trp Val Asp Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro
                405                 410                 415
Ala His Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr
            420                 425                 430
Cys Gly Val Gly
        435
```

<210> 57
<211> 1311
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetically generated oligonucleotide

<400> 57
atggccaagt acctggagct cgaagagagc ggggtcataa tgcaggcgtt ctactgggac 60
gtgccttcag gaggaatatg gtgggacaca atacggcaga agataccgga gtggtacgat 120
gccggaatct ccgcaatatg gattcctccc gcgagcaagg gtatgagcgg cggctattcg 180
atgggctacg acccctacga ttattttgac ctcggtgagt actaccagaa gggaacggtg 240
gaaacgaggt tcggctcaaa gcaggagctc ataaacatga taaacaccgc ccacgcctac 300
ggcatcaagg tcatcgcaga catagtaatc aaccaccgcg ccggaggaga ccttgagtgg 360
aacccccttcg tcaatgacta cacctggacg gacttctcga aggtcgcttc cggcaagtac 420
acggccaact acctcgactt ccaccccaac gaggtcaagt gctgtgacga gggcacattt 480
ggaggcttcc cagacatagc ccacgagaag agctgggacc agcactggct ctgggcgagc 540
gatgagagct acgccgccta cctaaggagc atcggcgttg atgcctggcg ctttgactac 600
gtgaagggct acggagcgtg ggtcgtcaag gactggctca actggtgggg tggctgggcc 660
gtcggggagt actgggacac aaacgttgat gcactgctca actgggccta ctcgagcgat 720
gcaaaagtct tcgacttccc gctctactac aagatggacg aggccttcga taacaacaac 780
attcccgccc tggtggacgc cctcagatac ggtcagacag tggtcagccg cgacccgttc 840
aaggctgtga cgtttgtagc caaccacgat accgatataa tctggaacaa gtaccttgct 900
tatgctttca tcctcaccta cgaaggccag cccgtcatat tctaccgcga ctacgaggag 960
tggctcaaca aggacaggtt gaacaacctc atatggatac acgaccacct cgcaggtgga 1020
agcactgaca tcgtttacta cgacaacgac gagctgatat tcgtgagaaa cggctacgga 1080
agcaagccgg gactgataac atacatcaac ctcgcctcaa gcaaagccgg aaggtgggtc 1140
tacgttccga agttcgcggg agcgtgcatc cacgagtaca ccggcaacct cggcggctgg 1200
gtggacaagt gggtggactc aagcgggtgg gtgtacctcg aggcccctgc ccacgacccg 1260
gccaacggct attacggcta ctccgtctgg agctattgcg gtgttggctg a 1311

<210> 58
<211> 436
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetically generated peptide

<400> 58
Met Ala Lys Tyr Leu Glu Leu Glu Glu Ser Gly Val Ile Met Gln Ala
 1   5   10   15
Phe Tyr Trp Asp Val Pro Ser Gly Gly Ile Trp Trp Asp Thr Ile Arg
            20   25   30
Gln Lys Ile Pro Glu Trp Tyr Asp Ala Gly Ile Ser Ala Ile Trp Ile
        35   40   45
Pro Pro Ala Ser Lys Gly Met Ser Gly Gly Tyr Ser Met Gly Tyr Asp
    50   55   60
Pro Tyr Asp Tyr Phe Asp Leu Gly Glu Tyr Tyr Gln Lys Gly Thr Val
65   70   75   80
Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Ile Asn Met Ile Asn Thr
                85   90   95
Ala His Ala Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His
            100   105   110
Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Asn Asp Tyr Thr
        115   120   125
Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr
    130   135   140
Leu Asp Phe His Pro Asn Glu Val Lys Cys Cys Asp Glu Gly Thr Phe
145   150   155   160
Gly Gly Phe Pro Asp Ile Ala His Glu Lys Ser Trp Asp Gln His Trp
                165   170   175
Leu Trp Ala Ser Asp Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly
            180   185   190
Val Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp Val
        195   200   205
Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr
    210   215   220
Trp Asp Thr Asn Val Asp Ala Leu Leu Asn Trp Ala Tyr Ser Ser Asp
225   230   235   240
Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Glu Ala Phe
                245   250   255
Asp Asn Asn Asn Ile Pro Ala Leu Val Asp Ala Leu Arg Tyr Gly Gln
            260   265   270
Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn

```
        275                 280                 285
His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Leu Ala Tyr Ala Phe Ile
    290                 295                 300
Leu Thr Tyr Glu Gly Gln Pro Val Ile Phe Tyr Arg Asp Tyr Glu Glu
305                 310                 315                 320
Trp Leu Asn Lys Asp Arg Leu Asn Asn Leu Ile Trp Ile His Asp His
                325                 330                 335
Leu Ala Gly Gly Ser Thr Asp Ile Val Tyr Tyr Asp Asn Asp Glu Leu
            340                 345                 350
Ile Phe Val Arg Asn Gly Tyr Gly Ser Lys Pro Gly Leu Ile Thr Tyr
        355                 360                 365
Ile Asn Leu Ala Ser Ser Lys Ala Gly Arg Trp Val Tyr Val Pro Lys
    370                 375                 380
Phe Ala Gly Ala Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp
385                 390                 395                 400
Val Asp Lys Trp Val Asp Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro
                405                 410                 415
Ala His Asp Pro Ala Asn Gly Tyr Tyr Gly Tyr Ser Val Trp Ser Tyr
            420                 425                 430
Cys Gly Val Gly
        435
```

<210> 59
<211> 1311
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetically generated oligonucleotide

<400> 59

```
atggccaagt acctggagct cgaagagggc ggggtcataa tgcaggcgtt ctactgggac      60
gtgccttcag gaggaatatg gtgggacaca atacggcaga agataccgga gtggtacgat     120
gccggaatct ccgcaatatg gattcctccc gcgagcaagg gtatgagcgg cggctattcg     180
atgggctacg acccctacga ttattttgac ctcggtgagt actaccagaa gggaacggtg     240
gaaacgaggt tcggctcaaa gcaggagctc ataaacatga taaacaccgc ccacgcctac     300
ggcatcaagg tcatcgcaga catagtaatc aaccaccgcg ccggaggaga ccttgagtgg     360
aacccccttcg tcaatgacta cacctggacg gacttctcga aggtcgcttc cggcaagtac     420
acggccaact acctcgactt ccacccgaac gagctccatg cgggcgattc cggaacattt     480
ggaggctatc ccgacatatg ccacgacaag agctgggacc agtactggct ctgggccagc     540
caggagagct acgcggcata tctcaggagc atcggcatcg atgcctggcg cttcgactac     600
gtcaagggct atgctccctg ggtcgtcaag gactggctga actggtgggg aggctgggcg     660
gttggagagt actgggacac caacgtcgac gctgttctca actgggcata ctcgagcggt     720
gccaaggtct ttgacttcgc cctctactac aagatggacg aggccttcga taacaacaac     780
attcccgccc tggtggacgc cctcagatac ggtcagacag tggtcagccg cgacccgttc     840
aaggctgtga cgtttgtagc caaccacgat accgatataa tttggaacaa gtacccggcc     900
tacgccttca tcctcaccta cgagggccag ccgacgatat tctaccgcga ctacgaggag     960
tggctcaaca aggacaggct caagaacctc atctggatac acgaccacct cgccggtgga    1020
agcactgaca tcgtttacta cgacaacgac gagctgatat tcgtgagaaa cggctacgga    1080
agcaagccgg gactgataac atacatcaac ctcgcgtcaa gcaaagccgg aaggtgggtt    1140
tatgtgccga agttcgcggg cgcgtgcatc cacgagtata ctggtaacct cggaggctgg    1200
gtagacaagt acgtctactc aagcggctgg gtctatctcg aagctccagc ttacgaccct    1260
gccaacgggc agtatggcta ctccgtgtgg agctattgcg gtgttgggtg a             1311
```

<210> 60
<211> 436
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetically generated peptide

<400> 60

```
Met Ala Lys Tyr Leu Glu Leu Glu Glu Gly Gly Val Ile Met Gln Ala
 1               5                  10                  15
Phe Tyr Trp Asp Val Pro Ser Gly Gly Ile Trp Trp Asp Thr Ile Arg
            20                  25                  30
Gln Lys Ile Pro Glu Trp Tyr Asp Ala Gly Ile Ser Ala Ile Trp Ile
```

```
        35                  40                  45
Pro Pro Ala Ser Lys Gly Met Ser Gly Gly Tyr Ser Met Gly Tyr Asp
    50                  55                  60
Pro Tyr Asp Tyr Phe Asp Leu Gly Glu Tyr Tyr Gln Lys Gly Thr Val
65                  70                  75                  80
Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Ile Asn Met Ile Asn Thr
                85                  90                  95
Ala His Ala Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His
            100                 105                 110
Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Asn Asp Tyr Thr
        115                 120                 125
Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr
    130                 135                 140
Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe
145                 150                 155                 160
Gly Gly Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp
                165                 170                 175
Leu Trp Ala Ser Gln Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly
            180                 185                 190
Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Ala Pro Trp Val
        195                 200                 205
Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr
    210                 215                 220
Trp Asp Thr Asn Val Asp Ala Val Leu Asn Trp Ala Tyr Ser Ser Gly
225                 230                 235                 240
Ala Lys Val Phe Asp Phe Ala Leu Tyr Tyr Lys Met Asp Glu Ala Phe
                245                 250                 255
Asp Asn Asn Asn Ile Pro Ala Leu Val Asp Ala Leu Arg Tyr Gly Gln
            260                 265                 270
Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn
        275                 280                 285
His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile
    290                 295                 300
Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu
305                 310                 315                 320
Trp Leu Asn Lys Asp Arg Leu Lys Asn Leu Ile Trp Ile His Asp His
                325                 330                 335
Leu Ala Gly Gly Ser Thr Asp Ile Val Tyr Tyr Asp Asn Asp Glu Leu
            340                 345                 350
Ile Phe Val Arg Asn Gly Tyr Gly Ser Lys Pro Gly Leu Ile Thr Tyr
        355                 360                 365
Ile Asn Leu Ala Ser Ser Lys Ala Gly Arg Trp Val Tyr Val Pro Lys
    370                 375                 380
Phe Ala Gly Ala Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp
385                 390                 395                 400
Val Asp Lys Tyr Val Tyr Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro
                405                 410                 415
Ala Tyr Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr
            420                 425                 430
Cys Gly Val Gly
        435
```

<210> 61
<211> 1311
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetically generated oligonucleotide

<400> 61
atggccaagt actccgagct gaaaaagggc ggggtcataa tgcaggcgtt ctactgggac 60
gtgccttcag gaggaatatg gtgggacaca atacggcaga agataccgga gtggtacgag 120
gcgggaatat ccgccatttg gattcctccc gcgagcaagg gtatgagcgg cggctattcg 180
atgggctacg acccctacga ttattttgac ctcggtgagt actaccagaa gggaacggtg 240
gaaacgaggt tcggctcaaa gcaggagctc ataaacatga taaacaccgc ccacgcctac 300
ggcatcaagg tcatcgcaga catagtaatc aaccaccgcg ccggaggaga ccttgagtgg 360
aacccttcg tcaatgacta cacctggacg gacttctcga aggtcgcttc cggcaagtac 420

acggccaact acctcaactt ccacccgaac gagctccatg cgggcgattc cggaacattt 480
ggaggctatc ccgacatatg ccacgacaag agctgggacc agtactggct ctgggccagc 540
caggagagct acgcggcata tctcaggagc atcggcatcg atgcctggcg cttcgactac 600
gtcaagggct acggagcgtg ggtcgtcaag gactggctgg actggtgggg aggctgggcc 660
gtcggggagt actgggacac aaacgttgat gcactgctca actgggccta ctcgagcgat 720
gcaaaagtct tcgacttccc gctctactac aagatggatg aggcctttga caacaaaaac 780
attccagcgc tcgtctctgc ccttcagaac ggccagactg ttgtctcccg cgacccgttc 840
aaggccgtaa cctttgtagc aaaccatgac accgatataa tctggaacaa gtatccagcc 900
tacgcgttca tcctcaccta cgagggccag ccgacaatat tctaccgcga ctacgaggag 960
tggctcaaca aggataagct caagaacctc atctggatac atgacaacct cgccggagga 1020
agcaccgaca tagtctacta cgataacgat gaactcatct tcgtcaggaa cggctacggg 1080
gacaagccgg ggcttataac ctacatcaac ctaggctcga gcaaggccgg aaggtgggtc 1140
tacgttccga agttcgcggg agcgtgcatc cacgagtaca ccggcaacct cggcggctgg 1200
gtggacaagt gggtggactc aagcgggtgg gtgtacctcg aggcccctgc ccacgacccg 1260
gccaacggct attacggcta ctccgtctgg agctactgcg gggtgggctg a 1311

<210> 62
<211> 436
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetically generated peptide

<400> 62
Met Ala Lys Tyr Ser Glu Leu Lys Lys Gly Gly Val Ile Met Gln Ala
 1   5   10   15
Phe Tyr Trp Asp Val Pro Ser Gly Gly Ile Trp Trp Asp Thr Ile Arg
            20   25   30
Gln Lys Ile Pro Glu Trp Tyr Glu Ala Gly Ile Ser Ala Ile Trp Ile
        35   40   45
Pro Pro Ala Ser Lys Gly Met Ser Gly Gly Tyr Ser Met Gly Tyr Asp
    50   55   60
Pro Tyr Asp Tyr Phe Asp Leu Gly Glu Tyr Tyr Gln Lys Gly Thr Val
65   70   75   80
Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Ile Asn Met Ile Asn Thr
                85   90   95
Ala His Ala Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His
            100   105   110
Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Asn Asp Tyr Thr
        115   120   125
Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr
    130   135   140
Leu Asn Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe
145   150   155   160
Gly Gly Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp
                165   170   175
Leu Trp Ala Ser Gln Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly
            180   185   190
Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp Val
        195   200   205
Val Lys Asp Trp Leu Asp Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr
    210   215   220
Trp Asp Thr Asn Val Asp Ala Leu Leu Asn Trp Ala Tyr Ser Ser Asp
225   230   235   240
Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Glu Ala Phe
                245   250   255
Asp Asn Lys Asn Ile Pro Ala Leu Val Ser Ala Leu Gln Asn Gly Gln
            260   265   270
Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn
        275   280   285
His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile
    290   295   300
Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu
305   310   315   320
Trp Leu Asn Lys Asp Lys Leu Lys Asn Leu Ile Trp Ile His Asp Asn
                325   330   335
Leu Ala Gly Gly Ser Thr Asp Ile Val Tyr Tyr Asp Asn Asp Glu Leu

340 345 350
Ile Phe Val Arg Asn Gly Tyr Gly Asp Lys Pro Gly Leu Ile Thr Tyr
355 360 365
Ile Asn Leu Gly Ser Ser Lys Ala Gly Arg Trp Val Tyr Val Pro Lys
370 375 380
Phe Ala Gly Ala Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp
385 390 395 400
Val Asp Lys Trp Val Asp Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro
405 410 415
Ala His Asp Pro Ala Asn Gly Tyr Tyr Gly Tyr Ser Val Trp Ser Tyr
420 425 430
Cys Gly Val Gly
435

<210> 63
<211> 1311
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetically generated oligonucleotide

<400> 63
atggccaagt acctggagct cgaagagggc ggggtcataa tgcaggcgtt ctactgggac 60
gtgccttcag gaggaatatg gtgggacaca atacggcaga agataccgga gtggtacgat 120
gccggaatct ccgcaatatg gattcccccg gcgagcaagg gcatgggcgg cgcctattcg 180
atgggctacg acccctacga cttctttgac ctcggtgagt acgaccagaa gggaacggta 240
gagacgcgct ttggctccaa gcaggagctc gtgaacatga taaacacggc ccatgcctac 300
ggcataaagg ccatagcgga catcgtcata aaccaccgcg caggcggaga cctcgagtgg 360
aacccgttcg ttggggacta cacctggacg gacttctcaa aggtggcctc gggcaaatat 420
actgccaact acctcgactt ccaccccaac gaggtcaagt gctgtgacga gggcacattt 480
ggaggcttcc cagacatagc ccacgagaag agctgggacc agcactggct ctgggcgagc 540
gatgagagct acgccgccta cctaaggagc atcggcgttg atgcctggcg ctttgactac 600
gtgaagggct acggagcgtg ggtcgtcaag gactggctca actggtgggg cggctgggcc 660
gttggcgagt actgggacac caacgttgat gcactcctca actgggccta ctcgagcggc 720
gccaaggtct tcgacttccc gctctactac aagatggacg cggcctttga caacaagaac 780
attcccgcac tcgtcgaggc cctcaagaac gggggcacag tcgtcagccg cgacccgttt 840
aaggccgtaa ccttcgttgc aaaccacgac accgatataa tctggaacaa gtatccagcc 900
tacgcgttca tcctcaccta cgagggccag ccgacaatat tctaccgcga ctacgaggag 960
tggctcaaca aggataagct caagaacctc atctggatac atgacaacct cgccggagga 1020
agcaccgaca tagtctacta cgataacgat gaactcatct tcgtcaggaa cggctacggg 1080
gacaagccgg ggcttataac ctacatcaac ctaggctgga gcaaggccgg aaggtgggtt 1140
tatgtgccga agttcgcggg cgcgtgcatc cacgagtata ctggtaacct cggaggctgg 1200
gtagacaagt acgtctactc aagcggctgg gtctatctcg aagctccagc ttacgaccct 1260
gccaacgggc agtatggcta ctccgtgtgg agctactgcg gggtggggtg a 1311

<210> 64
<211> 436
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetically generated peptide

<400> 64
Met Ala Lys Tyr Leu Glu Leu Glu Glu Gly Gly Val Ile Met Gln Ala
1 5 10 15
Phe Tyr Trp Asp Val Pro Ser Gly Gly Ile Trp Trp Asp Thr Ile Arg
20 25 30
Gln Lys Ile Pro Glu Trp Tyr Asp Ala Gly Ile Ser Ala Ile Trp Ile
35 40 45
Pro Pro Ala Ser Lys Gly Met Gly Gly Ala Tyr Ser Met Gly Tyr Asp
50 55 60
Pro Tyr Asp Phe Phe Asp Leu Gly Glu Tyr Asp Gln Lys Gly Thr Val
65 70 75 80
Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Val Asn Met Ile Asn Thr
85 90 95
Ala His Ala Tyr Gly Ile Lys Ala Ile Ala Asp Ile Val Ile Asn His

```
            100                 105                 110
Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Gly Asp Tyr Thr
        115                 120                 125
Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr
    130                 135                 140
Leu Asp Phe His Pro Asn Glu Val Lys Cys Cys Asp Glu Gly Thr Phe
145                 150                 155                 160
Gly Gly Phe Pro Asp Ile Ala His Glu Lys Ser Trp Asp Gln His Trp
                165                 170                 175
Leu Trp Ala Ser Asp Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly
            180                 185                 190
Val Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp Val
        195                 200                 205
Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr
    210                 215                 220
Trp Asp Thr Asn Val Asp Ala Leu Leu Asn Trp Ala Tyr Ser Ser Gly
225                 230                 235                 240
Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Ala Ala Phe
                245                 250                 255
Asp Asn Lys Asn Ile Pro Ala Leu Val Glu Ala Leu Lys Asn Gly Gly
            260                 265                 270
Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn
        275                 280                 285
His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile
    290                 295                 300
Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu
305                 310                 315                 320
Trp Leu Asn Lys Asp Lys Leu Lys Asn Leu Ile Trp Ile His Asp Asn
                325                 330                 335
Leu Ala Gly Gly Ser Thr Asp Ile Val Tyr Tyr Asp Asn Asp Glu Leu
            340                 345                 350
Ile Phe Val Arg Asn Gly Tyr Gly Asp Lys Pro Gly Leu Ile Thr Tyr
        355                 360                 365
Ile Asn Leu Gly Trp Ser Lys Ala Gly Arg Trp Val Tyr Val Pro Lys
    370                 375                 380
Phe Ala Gly Ala Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp
385                 390                 395                 400
Val Asp Lys Tyr Val Tyr Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro
                405                 410                 415
Ala Tyr Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr
            420                 425                 430
Cys Gly Val Gly
        435
```

<210> 65
<211> 1311
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetically generated oligonucleotide

<400> 65

```
atggccaagt actccgagct ggaagaaggc ggcgttataa tgcaggcctt ctactgggac      60
gtcccaggtg gaggaatctg gtggggcacc atcaggagca agataccgga gtggtacgag     120
gcgggaatat ccgccatttg gattcctccc gcgagcaagg gtatgagcgg cggctattcg     180
atgggctacg acccctacga ttattttgac ctcggtgagt actaccagaa gggaacggtg     240
gaaacgaggt tcggctcaaa gcaggagctc ataaacatga taaacaccgc ccacgcctat     300
ggcatgaagg taatagccga tatagtcatc aaccaccgcg ccggcggtga cctggagtgg     360
aacccttcg tgaacgacta tacctggacc gacttctcaa aggtcgcgtc gggtaaatac     420
acggccaact acctcgactt ccacccgaac gagctccatg cgggcgattc cggaacattt     480
ggaggctatc ccgacatatg ccacgacaag agctgggacc agtactggct ctgggccagc     540
caggagagct acgcggcata tctcaggagc atcggcatcg atgcctggcg cttcgactac     600
gtcaagggct atgctccctg ggtcgtcaag gactggctga actggtgggg aggctgggcg     660
gttggagagt actgggacac caacgtcgac gctgttctca actgggcata ctcgagcggt     720
gccaaggtct ttgacttcgc cctctactac aagatggacg aggccttcga taacaacaac     780
attcccgccc tggtggacgc cctcagatac ggtcagacag tggtcagccg cgacccgttc     840
aaggctgtga cgtttgtagc caaccacgat accgatataa tttggaacaa gtacccggcc     900
```

tacgccttca tcctcaccta cgagggccag ccgacgatat tctaccgcga ctacgaggag 960
tggctcaaca aggacaggct caagaacctc atctggatac acgaccacct cgccggtgga 1020
agcacgagca tagtttacta cgacagcgac gagatgatct tcgtgaggaa cggctatgga 1080
agcaagcctg gccttataac ttacatcaac ctcggctcga gcaaggttgg aaggtgggtt 1140
tacgttccga agttcgcagg ctcgtgcata cacgagtaca ccggcaatct cggcggctgg 1200
gtggacaagt gggtggactc aagcggctgg gtctacctcg aggctcctgc ccacgacccg 1260
gccaacggcc agtacggcta ctccgtctgg agctattgcg gtgttggctg a 1311

<210> 66
<211> 436
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetically generated peptide

<400> 66

```
Met Ala Lys Tyr Ser Glu Leu Glu Glu Gly Gly Val Ile Met Gln Ala
 1               5                  10                  15
Phe Tyr Trp Asp Val Pro Gly Gly Gly Ile Trp Trp Gly Thr Ile Arg
            20                  25                  30
Ser Lys Ile Pro Glu Trp Tyr Glu Ala Gly Ile Ser Ala Ile Trp Ile
        35                  40                  45
Pro Pro Ala Ser Lys Gly Met Ser Gly Gly Tyr Ser Met Gly Tyr Asp
    50                  55                  60
Pro Tyr Asp Tyr Phe Asp Leu Gly Glu Tyr Tyr Gln Lys Gly Thr Val
65                  70                  75                  80
Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Ile Asn Met Ile Asn Thr
                85                  90                  95
Ala His Ala Tyr Gly Met Lys Val Ile Ala Asp Ile Val Ile Asn His
            100                 105                 110
Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Asn Asp Tyr Thr
        115                 120                 125
Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr
    130                 135                 140
Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe
145                 150                 155                 160
Gly Gly Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp
                165                 170                 175
Leu Trp Ala Ser Gln Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly
            180                 185                 190
Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Ala Pro Trp Val
        195                 200                 205
Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr
    210                 215                 220
Trp Asp Thr Asn Val Asp Ala Val Leu Asn Trp Ala Tyr Ser Ser Gly
225                 230                 235                 240
Ala Lys Val Phe Asp Phe Ala Leu Tyr Tyr Lys Met Asp Glu Ala Phe
                245                 250                 255
Asp Asn Asn Asn Ile Pro Ala Leu Val Asp Ala Leu Arg Tyr Gly Gln
            260                 265                 270
Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn
        275                 280                 285
His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile
    290                 295                 300
Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu
305                 310                 315                 320
Trp Leu Asn Lys Asp Arg Leu Lys Asn Leu Ile Trp Ile His Asp His
                325                 330                 335
Leu Ala Gly Gly Ser Thr Ser Ile Val Tyr Tyr Asp Ser Asp Glu Met
            340                 345                 350
Ile Phe Val Arg Asn Gly Tyr Gly Ser Lys Pro Gly Leu Ile Thr Tyr
        355                 360                 365
Ile Asn Leu Gly Ser Ser Lys Val Gly Arg Trp Val Tyr Val Pro Lys
    370                 375                 380
Phe Ala Gly Ser Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp
385                 390                 395                 400
Val Asp Lys Trp Val Asp Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro
```

405 410 415
Ala His Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr
420 425 430
Cys Gly Val Gly
435

<210> 67
<211> 1311
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetically generated oligonucleotide

<400> 67
atggccaagt acctggagct cgaagagggc ggggtcataa tgcaggcgtt ctactgggac 60
gtgccttcgg gaggaatatg gtgggacaca atacggcaga agataccgga gtggtacgat 120
gccggaatct ccgcaatatg gattcctccc gcgagcaagg gtatgagcgg cggctattcg 180
atgggctacg acccctacga ttattttgac ctcggtgagt actaccagaa gggaacggtg 240
gaaacgaggt tcggctcaaa gcaggagctc ataaacatga taaacacggc ccatgcctac 300
ggcataaagg tcatagcgga catcgtcata aaccaccgcg caggcggaga cctcgagtgg 360
aacccgttcg ttggggacta cacctggacg gacttctcaa aggtggcctc gggcaaatat 420
actgccaact acctcgactt ccaccccaac gaggtcaagt gctgtgacga gggcacattt 480
ggaggcttcc cagacatagc ccacgagaag agctgggacc agcactggct ctgggcgagc 540
gatgagagct acgccgccta cctaaggagc atcggcgttg atgcctggcg cttcgactac 600
gtcaagggct acggagcgtg ggtcgtcaag gactggctgg actggtgggg aggctgggcc 660
gtcggggagt actgggacac aaacgttgat gcactgctca actgggccta ctcgagcgat 720
gcaaaagtct tcgacttccc gctctactac aagatggacg aggccttcga taacaacaac 780
attcccgccc tggtggacgc cctcagatac ggtcagacag tggtcagccg cgacccgttc 840
aaggctgtga cgtttgtagc caaccacgat accgatataa tctggaacaa gtatccagcc 900
tacgcgttca tcctcaccta cgagggccag ccgacaatat tctaccgcga ctacgaggag 960
tggctcaaca aggataagct caagaacctc atctggatac atgacaacct cgccggagga 1020
agcacgagca tagtttacta cgacagcgac gagatgatct tcgtgaggaa cggctatgga 1080
agcaagcctg gccttataac ttacatcaac ctcggctcga gcaaggttgg aaggtgggtc 1140
tacgttccga agttcgcggg agcgtgcatc cacgagtaca ccggcaacct cggcggctgg 1200
gtggacaagt gggtggactc aagcgggtgg gtgtacctcg aggcccctgc ccacgacccg 1260
gccaacggct attacggcta ctccgtctgg agctactgcg tggtgggctg a 1311

<210> 68
<211> 436
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetically generated peptide

<400> 68
Met Ala Lys Tyr Leu Glu Leu Glu Glu Gly Gly Val Ile Met Gln Ala
1 5 10 15
Phe Tyr Trp Asp Val Pro Ser Gly Gly Ile Trp Trp Asp Thr Ile Arg
20 25 30
Gln Lys Ile Pro Glu Trp Tyr Asp Ala Gly Ile Ser Ala Ile Trp Ile
35 40 45
Pro Pro Ala Ser Lys Gly Met Ser Gly Gly Tyr Ser Met Gly Tyr Asp
50 55 60
Pro Tyr Asp Tyr Phe Asp Leu Gly Glu Tyr Tyr Gln Lys Gly Thr Val
65 70 75 80
Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Ile Asn Met Ile Asn Thr
85 90 95
Ala His Ala Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His
100 105 110
Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Gly Asp Tyr Thr
115 120 125
Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr
130 135 140
Leu Asp Phe His Pro Asn Glu Val Lys Cys Cys Asp Glu Gly Thr Phe
145 150 155 160
Gly Gly Phe Pro Asp Ile Ala His Glu Lys Ser Trp Asp Gln His Trp

```
                165                 170                 175
Leu Trp Ala Ser Asp Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly
            180                 185                 190
Val Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp Val
        195                 200                 205
Val Lys Asp Trp Leu Asp Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr
    210                 215                 220
Trp Asp Thr Asn Val Asp Ala Leu Leu Asn Trp Ala Tyr Ser Ser Asp
225                 230                 235                 240
Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Glu Ala Phe
                245                 250                 255
Asp Asn Asn Asn Ile Pro Ala Leu Val Asp Ala Leu Arg Tyr Gly Gln
            260                 265                 270
Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn
        275                 280                 285
His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile
    290                 295                 300
Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu
305                 310                 315                 320
Trp Leu Asn Lys Asp Lys Leu Lys Asn Leu Ile Trp Ile His Asp Asn
                325                 330                 335
Leu Ala Gly Gly Ser Thr Ser Ile Val Tyr Tyr Asp Ser Asp Glu Met
            340                 345                 350
Ile Phe Val Arg Asn Gly Tyr Gly Ser Lys Pro Gly Leu Ile Thr Tyr
        355                 360                 365
Ile Asn Leu Gly Ser Ser Lys Val Gly Arg Trp Val Tyr Val Pro Lys
    370                 375                 380
Phe Ala Gly Ala Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp
385                 390                 395                 400
Val Asp Lys Trp Val Asp Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro
                405                 410                 415
Ala His Asp Pro Ala Asn Gly Tyr Tyr Gly Tyr Ser Val Trp Ser Tyr
            420                 425                 430
Cys Val Val Gly
        435
```

<210> 69
<211> 1542
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 69

```
atgttgaaaa ggattacggt agtctgttta ttgtttattt tgctttttcc taatatatat      60
gagggaaata aggcagaagc agcaacagtg aacaatggaa cattaatgca gtattttgag     120
tggtacgctc cgaatgatgg gaatcattgg aatcgtttgc gttccgatgc tgaaagttta     180
gctcataaag gaatcacatc tgtatggata ccacctgcat ataaagggac ttcgcaaaat     240
gatgtagggt atggggccta tgatttatat gatttagggg agttcaatca aaaaggaacg     300
gtgcggacga aatatgggac aaaagcacag ttgaaatctg caattgacgc tttacataag     360
caaaacatcg acgtatacgg tgatgtagtt atgaatcata aaggtggggc tgattatact     420
gaaaccgtaa cagctgttga ggtagaccgt aacaatcgaa atattgaagt atcaggtgat     480
tatcaaatta gtgcatggac ggggtttaat tttccagggc gcggagatgc ttattctaat     540
ttcaaatgga aatggtatca ttttgacgga acggattggg atgaaggaag gaaattaaat     600
cgaatttata aatttagggg tgtagataaa gcgtgggatt gggaagtgtc tagcgaaaat     660
ggaaattatg attatttgat gtatgcagat cttgattttg atcatcctga tgttgcgaat     720
gagatgaaaa attggggaac atggtatgcg aatgaattaa atttagatgg ctttcgtttg     780
gacgctgtta aacatattga tcatgaatat ttacgcgatt gggtaaatca tgccagacag     840
caaacgggga aagaaatgtt tacagtagct gaatattggc aaaatgatgt tcaggcttta     900
aacaattatt tagcgaaagt caattataat caatctgtgt ttgatgcacc gcttcattac     960
aattttcatt atgcttcaac aggaaatggg aattatgata tgagaaatat tttaaatgga    1020
acagtaatga aaaatcaccc tgcactcgca gttactctcg ttgagaatca tgattctcag    1080
cctgggcagt cattggaatc tgtagtaagt ccgtggttta agccgctggc atatgcattt    1140
attttaactc gtgcagaggg ctatccttca gttttctatg gtgattacta tgggacaagc    1200
ggaaatagta gttatgaaat tccagcgtta aaagataaaa ttgatccaat tttgacggca    1260
cgaaaaaact ttgcatatgg tacgcagcgt gattatttag accatccaga tgtgattggc    1320
tggacaagag aaggcgatgg tgtacatgct aattctggtt tagcgacatt actctcggac    1380
```

ggaccaggag gatcaaagtg gatggatgtt ggaaagaata acgctgggga agtatggtac 1440
gatattacgg gtaatcaaac aaatactgta acaattaata aggacggatg ggggcagttc 1500
tatgtaagtg gcggctcagt ttccatatat gttcagcggt aa 1542

<210> 70
<211> 513
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 70

```
Met Leu Lys Arg Ile Thr Val Val Cys Leu Leu Phe Ile Leu Leu Phe
 1               5                  10                  15
Pro Asn Ile Tyr Glu Gly Asn Lys Ala Glu Ala Ala Thr Val Asn Asn
            20                  25                  30
Gly Thr Leu Met Gln Tyr Phe Glu Trp Tyr Ala Pro Asn Asp Gly Asn
        35                  40                  45
His Trp Asn Arg Leu Arg Ser Asp Ala Glu Ser Leu Ala His Lys Gly
    50                  55                  60
Ile Thr Ser Val Trp Ile Pro Pro Ala Tyr Lys Gly Thr Ser Gln Asn
65                  70                  75                  80
Asp Val Gly Tyr Gly Ala Tyr Asp Leu Tyr Asp Leu Gly Glu Phe Asn
                85                  90                  95
Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly Thr Lys Ala Gln Leu Lys
            100                 105                 110
Ser Ala Ile Asp Ala Leu His Lys Gln Asn Ile Asp Val Tyr Gly Asp
        115                 120                 125
Val Val Met Asn His Lys Gly Gly Ala Asp Tyr Thr Glu Thr Val Thr
    130                 135                 140
Ala Val Glu Val Asp Arg Asn Asn Arg Asn Ile Glu Val Ser Gly Asp
145                 150                 155                 160
Tyr Gln Ile Ser Ala Trp Thr Gly Phe Asn Phe Pro Gly Arg Gly Asp
                165                 170                 175
Ala Tyr Ser Asn Phe Lys Trp Lys Trp Tyr His Phe Asp Gly Thr Asp
            180                 185                 190
Trp Asp Glu Gly Arg Lys Leu Asn Arg Ile Tyr Lys Phe Arg Gly Val
        195                 200                 205
Asp Lys Ala Trp Asp Trp Glu Val Ser Ser Glu Asn Gly Asn Tyr Asp
    210                 215                 220
Tyr Leu Met Tyr Ala Asp Leu Asp Phe Asp His Pro Asp Val Ala Asn
225                 230                 235                 240
Glu Met Lys Asn Trp Gly Thr Trp Tyr Ala Asn Glu Leu Asn Leu Asp
                245                 250                 255
Gly Phe Arg Leu Asp Ala Val Lys His Ile Asp His Glu Tyr Leu Arg
            260                 265                 270
Asp Trp Val Asn His Ala Arg Gln Gln Thr Gly Lys Glu Met Phe Thr
        275                 280                 285
Val Ala Glu Tyr Trp Gln Asn Asp Val Gln Ala Leu Asn Asn Tyr Leu
    290                 295                 300
Ala Lys Val Asn Tyr Asn Gln Ser Val Phe Asp Ala Pro Leu His Tyr
305                 310                 315                 320
Asn Phe His Tyr Ala Ser Thr Gly Asn Gly Asn Tyr Asp Met Arg Asn
                325                 330                 335
Ile Leu Asn Gly Thr Val Met Lys Asn His Pro Ala Leu Ala Val Thr
            340                 345                 350
Leu Val Glu Asn His Asp Ser Gln Pro Gly Gln Ser Leu Glu Ser Val
        355                 360                 365
Val Ser Pro Trp Phe Lys Pro Leu Ala Tyr Ala Phe Ile Leu Thr Arg
    370                 375                 380
Ala Glu Gly Tyr Pro Ser Val Phe Tyr Gly Asp Tyr Tyr Gly Thr Ser
385                 390                 395                 400
Gly Asn Ser Ser Tyr Glu Ile Pro Ala Leu Lys Asp Lys Ile Asp Pro
                405                 410                 415
Ile Leu Thr Ala Arg Lys Asn Phe Ala Tyr Gly Thr Gln Arg Asp Tyr
            420                 425                 430
Leu Asp His Pro Asp Val Ile Gly Trp Thr Arg Glu Gly Asp Gly Val
```

```
        435                 440                 445
His Ala Asn Ser Gly Leu Ala Thr Leu Leu Ser Asp Gly Pro Gly Gly
    450                 455                 460
Ser Lys Trp Met Asp Val Gly Lys Asn Asn Ala Gly Glu Val Trp Tyr
465                 470                 475                 480
Asp Ile Thr Gly Asn Gln Thr Asn Thr Val Thr Ile Asn Lys Asp Gly
                485                 490                 495
Trp Gly Gln Phe Tyr Val Ser Gly Gly Ser Val Ser Ile Tyr Val Gln
            500                 505                 510
Arg
```

<210> 71
<211> 1311
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetically generated oligonucleotide

<400> 71

```
atggccaagt acctggagct cgaagagggc ggggtcataa tgcaggcgtt ctactgggac      60
gtgccttcag gaggaatatg gtgggacaca atacggcaga agataccgga gtggtacgat     120
gccggaatct ccgcaatatg gattcccccg gcgagcaagg gcatgggcgg cgcctattcg     180
atgggctacg acccctacga cttctttgac ctcggtgagt acgaccagaa gggaacggta     240
gagacgcgct ttggctccaa gcaggagctc gtgaacatga taaacacggc ccatgcctac     300
ggcataaagg tcatagcgga catcgtcata aaccaccgcg caggcggaga cctcgagtgg     360
aacccgttcg ttggggacta cacctggacg gacttctcaa aggtagcctc gggcaaatat     420
actgccaact acctcgactt ccacccgaac gagctccatg cgggcgattc cggaacattt     480
ggaggctatc ccgacatatg ccacgacaag agctgggacc agtactggct ctgggccagc     540
caggagagct acgcggcata tctcaggagc atcggcatcg atgcctggcg cttcgactac     600
gtcaagggct atgctccctg ggtcgtcaag gactggctga actggtgggg aggctgggcg     660
gttggagagt actgggacac caacgtcgac gctgttctca actgggcata ctcgagcggt     720
gccaaggtct ttgacttcgc cctctactac aagatggatg aggcctttga caacaaaaac     780
attccagcgc tcgtctctgc ccttcagaac ggccagactg ttgtctcccg cgacccgttc     840
aaggccgtaa cctttgtagc aaaccacgac accgatataa tctggaacaa gtatccagcc     900
tacgcgttca tcctcaccta cgagggccag ccgacaatat tctaccgcga ctacgaggag     960
tggctcaaca aggataagct caagaacctc atctggatac atgacaacct cgccggagga    1020
agcactgaca tcgtttacta cgacaacgac gagctgatat tcgtgagaaa cggctacgga    1080
agcaagccgg gactgataac atacatcaac ctcgcctcaa gcaaagccgg aaggtgggtt    1140
tatgtgccga agttcgcggg cgcgtgcatc cacgagtata ctggtaacct cggaggctgg    1200
gtagacaagt acgtctactc aagcggctgg gtctatctcg aagctccagc ttacgaccct    1260
gccaacgggc agtatggcta ctccgtgtgg agctactgcg gggtgggctg a             1311
```

<210> 72
<211> 436
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetically generated peptide

<400> 72

```
Met Ala Lys Tyr Leu Glu Leu Glu Glu Gly Gly Val Ile Met Gln Ala
 1               5                  10                  15
Phe Tyr Trp Asp Val Pro Ser Gly Gly Ile Trp Trp Asp Thr Ile Arg
            20                  25                  30
Gln Lys Ile Pro Glu Trp Tyr Asp Ala Gly Ile Ser Ala Ile Trp Ile
        35                  40                  45
Pro Pro Ala Ser Lys Gly Met Gly Gly Ala Tyr Ser Met Gly Tyr Asp
    50                  55                  60
Pro Tyr Asp Phe Phe Asp Leu Gly Glu Tyr Asp Gln Lys Gly Thr Val
65                  70                  75                  80
Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Val Asn Met Ile Asn Thr
                85                  90                  95
Ala His Ala Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His
            100                 105                 110
Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Gly Asp Tyr Thr
```

```
        115                 120                 125
Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr
    130                 135                 140
Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe
145                 150                 155                 160
Gly Gly Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp
                165                 170                 175
Leu Trp Ala Ser Gln Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly
            180                 185                 190
Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Ala Pro Trp Val
        195                 200                 205
Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr
    210                 215                 220
Trp Asp Thr Asn Val Asp Ala Val Leu Asn Trp Ala Tyr Ser Ser Gly
225                 230                 235                 240
Ala Lys Val Phe Asp Phe Ala Leu Tyr Tyr Lys Met Asp Glu Ala Phe
                245                 250                 255
Asp Asn Lys Asn Ile Pro Ala Leu Val Ser Ala Leu Gln Asn Gly Gln
            260                 265                 270
Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn
        275                 280                 285
His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile
    290                 295                 300
Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu
305                 310                 315                 320
Trp Leu Asn Lys Asp Lys Leu Lys Asn Leu Ile Trp Ile His Asp Asn
                325                 330                 335
Leu Ala Gly Gly Ser Thr Asp Ile Val Tyr Tyr Asp Asn Asp Glu Leu
            340                 345                 350
Ile Phe Val Arg Asn Gly Tyr Gly Ser Lys Pro Gly Leu Ile Thr Tyr
        355                 360                 365
Ile Asn Leu Ala Ser Ser Lys Ala Gly Arg Trp Val Tyr Val Pro Lys
    370                 375                 380
Phe Ala Gly Ala Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp
385                 390                 395                 400
Val Asp Lys Tyr Val Tyr Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro
                405                 410                 415
Ala Tyr Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr
            420                 425                 430
Cys Gly Val Gly
        435
```

<210> 73
<211> 1299
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 73

```
atggctctgg aagagggcgg gcttataatg caggccttct actgggacgt cccaggtgga      60
ggaatctggt gggacaccat agcccagaag atacccgact gggcgagcgc cgggatttcg     120
gcaatatgga ttcctcccgc gagtaagggc atgagcggcg gctattcgat gggctacgac     180
ccctacgatt tcttcgacct cggtgagtac taccagaagg gaagcgttga gacccgcttc     240
ggatcaaaag aggagcttgt gaacatgata aacaccgccc atgctcacaa catgaaggtc     300
atagcggaca tagtcatcaa ccaccgcgcc ggcggcgacc tggagtggaa tcctttcacc     360
aacagctaca cctggaccga tttctcgaag gtcgcgtcgg gcaagtacac ggccaactac     420
ctcgacttcc acccgaacga gcttcacgcg ggcgattccg gaacatttgg aggctatccc     480
gacatatgcc acgacaagag ctgggaccag cactggctct gggccagcaa cgaaagctac     540
gccgcctacc tccggagcat cggcatcgac gcctggcgct tcgactacgt caagggctac     600
gctccctggg tcgttaagaa ctggctgaac cggtggggcg gctgggcggt tggagagtac     660
tgggacacca acgtcgatgc actcctgagc tgggcctacg acagcggtgc taaagtcttc     720
gacttcccgc tctactacaa gatggacgag gccttcgata acaacaacat ccccgccctc     780
gtggacgccc tcaagaacgg aggcacggtc gtcagccgcg acccgttcaa agccgtgacc     840
ttcgttgcca accacgatac caacataatc tggaacaagt atccggccta cgccttcatc     900
ctcacctatg agggacagcc ggcaatattc taccgcgact acgaggagtg gctcaacaag     960
gacaggctca ggaacctcat ctggatacac gaccacctcg cgggaggaag cacagacatc    1020
```

atctactacg acagcgacga gcttatcttc gtgagaaacg gctacgggga caagccggga 1080
ctgataacct acatcaacct cggctcaagc aaggccggaa ggtgggtcta cgttccgaag 1140
ttcgcaggct cgtgcataca cgagtacacc ggcaacctcg gcggctggat tgacaagtgg 1200
gttgactcaa gcggtcgggt ctaccttgag gcccccgccc acgacccggc caacggccag 1260
tacggctact ccgtatggag ctactgcggt gttgggtga 1299

<210> 74
<211> 432
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 74
Met Ala Leu Glu Glu Gly Gly Leu Ile Met Gln Ala Phe Tyr Trp Asp
 1   5   10   15
Val Pro Gly Gly Gly Ile Trp Trp Asp Thr Ile Ala Gln Lys Ile Pro
            20   25   30
Asp Trp Ala Ser Ala Gly Ile Ser Ala Ile Trp Ile Pro Pro Ala Ser
        35   40   45
Lys Gly Met Ser Gly Gly Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Phe
    50   55   60
Phe Asp Leu Gly Glu Tyr Tyr Gln Lys Gly Ser Val Glu Thr Arg Phe
65   70   75   80
Gly Ser Lys Glu Glu Leu Val Asn Met Ile Asn Thr Ala His Ala His
                85   90   95
Asn Met Lys Val Ile Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly
            100   105   110
Asp Leu Glu Trp Asn Pro Phe Thr Asn Ser Tyr Thr Trp Thr Asp Phe
        115   120   125
Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr Leu Asp Phe His
    130   135   140
Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe Gly Gly Tyr Pro
145   150   155   160
Asp Ile Cys His Asp Lys Ser Trp Asp Gln His Trp Leu Trp Ala Ser
                165   170   175
Asn Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly Ile Asp Ala Trp
            180   185   190
Arg Phe Asp Tyr Val Lys Gly Tyr Ala Pro Trp Val Val Lys Asn Trp
        195   200   205
Leu Asn Arg Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn
    210   215   220
Val Asp Ala Leu Leu Ser Trp Ala Tyr Asp Ser Gly Ala Lys Val Phe
225   230   235   240
Asp Phe Pro Leu Tyr Tyr Lys Met Asp Glu Ala Phe Asp Asn Asn Asn
                245   250   255
Ile Pro Ala Leu Val Asp Ala Leu Lys Asn Gly Gly Thr Val Val Ser
            260   265   270
Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Asn
        275   280   285
Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile Leu Thr Tyr Glu
    290   295   300
Gly Gln Pro Ala Ile Phe Tyr Arg Asp Tyr Glu Glu Trp Leu Asn Lys
305   310   315   320
Asp Arg Leu Arg Asn Leu Ile Trp Ile His Asp His Leu Ala Gly Gly
                325   330   335
Ser Thr Asp Ile Ile Tyr Tyr Asp Ser Asp Glu Leu Ile Phe Val Arg
            340   345   350
Asn Gly Tyr Gly Asp Lys Pro Gly Leu Ile Thr Tyr Ile Asn Leu Gly
        355   360   365
Ser Ser Lys Ala Gly Arg Trp Val Tyr Val Pro Lys Phe Ala Gly Ser
    370   375   380
Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp Ile Asp Lys Trp
385   390   395   400
Val Asp Ser Ser Gly Arg Val Tyr Leu Glu Ala Pro Ala His Asp Pro
                405   410   415
Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr Cys Gly Val Gly

420 425 430

<210> 75
<211> 1299
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 75
atggctctgg aagagggcgg gcttataatg caggcattct actgggacgt ccccatggga 60
ggaatctggt gggacacgat agcccagaag atacccgact gggcaagcgc cgggatttcg 120
gcgatatgga ttccccccgc gagcaagggt atgagcggcg gctattcgat gggctacgac 180
ccctacgatt attttgacct cggtgagtac taccagaagg gaacggtgga aacaagattc 240
ggctcaaagc aggagctcat aaacatgata aacaccgccc acgcctatgg catgaaggta 300
atagccgata tagtcatcaa ccaccgcgcc ggcggcgatc tggagtggaa ccccttcgtg 360
aacgactata cctggaccga cttctcgaag gtcgcgtcgg gtaaatacac ggccaactac 420
ctcgacttcc acccgaacga gctccacgcg ggcgattccg gaacatttgg aggctatccc 480
gacatatgcc acgacaagag ctgggaccag tactggctct gggccagcca ggagagctac 540
gcggcctatc tcaggagcat cggcatcgac gcctggcgct tcgactacgt caagggctat 600
gctccctggg tcgtcaggga ctggctgaac tggtggggag gctgggcagt tggagagtac 660
tgggacacca acgtcgacgc tgttctcaac tgggcatact cgagcggtgc caaggtcttt 720
gacttcgccc tctactacaa gatggacgag gccttcgata acaacaacat tcccgccctg 780
gtggacgccc tcagatacgg ccagacagtg gtcagccgcg acccgttcaa ggctgtgacg 840
tttgtagcca accacgatac cgacataatc tggaacaagt atccagccta cgcgttcatc 900
ctcacctacg agggccagcc gacaatattc taccgcgact acgaggagtg gctcaacaag 960
gacaagctca agaacctcat ctggatacat gacaacctcg ccggagggag cactgacatc 1020
gtttactacg acaacgacga gctgatattc gtgagaaacg gctacggaag caagccggga 1080
ctgataacat acatcaacct cggctcaagc aaagccggaa ggtgggttta cgttccgaag 1140
ttcgcaggct cgtgcataca cgagtacacc ggcaacctcg gcggctgggt ggacaagtgg 1200
gtggactcaa gcggctgggt ttacctcgag gctcctgccc acgacccggc caacggccag 1260
tacggctact ccgtttggag ctattgcggt gttgggtga 1299

<210> 76
<211> 432
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 76
Met Ala Leu Glu Glu Gly Gly Leu Ile Met Gln Ala Phe Tyr Trp Asp
1 5 10 15
Val Pro Met Gly Gly Ile Trp Trp Asp Thr Ile Ala Gln Lys Ile Pro
20 25 30
Asp Trp Ala Ser Ala Gly Ile Ser Ala Ile Trp Ile Pro Pro Ala Ser
35 40 45
Lys Gly Met Ser Gly Gly Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Tyr
50 55 60
Phe Asp Leu Gly Glu Tyr Tyr Gln Lys Gly Thr Val Glu Thr Arg Phe
65 70 75 80
Gly Ser Lys Gln Glu Leu Ile Asn Met Ile Asn Thr Ala His Ala Tyr
85 90 95
Gly Met Lys Val Ile Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly
100 105 110
Asp Leu Glu Trp Asn Pro Phe Val Asn Asp Tyr Thr Trp Thr Asp Phe
115 120 125
Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr Leu Asp Phe His
130 135 140
Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe Gly Gly Tyr Pro
145 150 155 160
Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp Leu Trp Ala Ser
165 170 175
Gln Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly Ile Asp Ala Trp
180 185 190
Arg Phe Asp Tyr Val Lys Gly Tyr Ala Pro Trp Val Val Arg Asp Trp

```
        195                 200                 205
Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn
    210                 215                 220
Val Asp Ala Val Leu Asn Trp Ala Tyr Ser Ser Gly Ala Lys Val Phe
225                 230                 235                 240
Asp Phe Ala Leu Tyr Tyr Lys Met Asp Glu Ala Phe Asp Asn Asn Asn
                245                 250                 255
Ile Pro Ala Leu Val Asp Ala Leu Arg Tyr Gly Gln Thr Val Val Ser
            260                 265                 270
Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Asp
        275                 280                 285
Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile Leu Thr Tyr Glu
    290                 295                 300
Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu Trp Leu Asn Lys
305                 310                 315                 320
Asp Lys Leu Lys Asn Leu Ile Trp Ile His Asp Asn Leu Ala Gly Gly
                325                 330                 335
Ser Thr Asp Ile Val Tyr Tyr Asp Asn Asp Glu Leu Ile Phe Val Arg
            340                 345                 350
Asn Gly Tyr Gly Ser Lys Pro Gly Leu Ile Thr Tyr Ile Asn Leu Gly
        355                 360                 365
Ser Ser Lys Ala Gly Arg Trp Val Tyr Val Pro Lys Phe Ala Gly Ser
    370                 375                 380
Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp Val Asp Lys Trp
385                 390                 395                 400
Val Asp Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro Ala His Asp Pro
                405                 410                 415
Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr Cys Gly Val Gly
            420                 425                 430
```

<210> 77
<211> 1299
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 77

```
atggctctgg aagagggcgg gctcataatg caggccttct actgggacgt ccccatggga      60
ggaatctggt gggacacgat agcccagaag atacccgact gggcaagcgc cgggatttcg     120
gcgatatgga tccctcccgc gagcaagggt atgagcggcg gctattcgat gggctacgac     180
ccctacgatt attttgacct cggtgagtac taccagaagg gaacggtgga aacgaggttc     240
ggctcaaagc aggagctcat aaacatgata aacaccgccc acgcctatgg catgaaggta     300
atagccgata tagtcatcaa ccaccgcgcc ggcggtgacc tggagtggaa ccccttcgtg     360
aacgactata cctggaccga cttctcaaag gtcgcgtcgg gtaaatacac ggccaactac     420
ctcgacttcc acccgaacga gctccatgcg ggcgattccg gaacatttgg aggctatccc     480
gacatatgcc acgacaagag ctgggaccag tactggctct gggccagcca ggagagctac     540
gcggcatatc tcaggagcat cggcatcgat gcctggcgct tcgactacgt caagggctat     600
gctccctggg tcgtcaagga ctggctgaac tggtggggag gctgggcggt tggagagtac     660
tgggacacca acgtcgacgc tgttctcaac tgggcatact cgagcggtgc caaggtcttt     720
gacttcgccc tctactacaa gatggacgag gccttcgata acaacaacat tcccgccctg     780
gtggacgccc tcagatacgg tcagacagtg gtcagccgcg acccgttcaa ggctgtgacg     840
tttgtagcca accacgatac cgacataatc tggaacaagt atccagccta cgcgttcatc     900
ctcacctacg agggccagcc gacaatattc taccgcgact acgaggagtg gctcaacaag     960
gataagctca agaacctcat ctggatacat gacaacctcg ccggagggag cactgacatc    1020
gtttactacg acaacgacga gctgatattc gtgagaaacg gctacggaag caagccggga    1080
ctgataacat acatcaacct cgcctcaagc aaagccggaa ggtgggttta cgttccgaag    1140
ttcgcaggct cgtgcataca cgagtacacc ggcaatctcg gcggctgggt ggacaagtgg    1200
gtggactcaa gcggctgggt ctacctcgag gctcctgccc acgacccggc caacggccag    1260
tacggctact ccgtctggag ctactgcggt gttgggtga                           1299
```

<210> 78
<211> 432
<212> PRT
<213> Unknown

<220>

<223> Obtained from an environmental sample.

<400> 78

```
Met Ala Leu Glu Glu Gly Gly Leu Ile Met Gln Ala Phe Tyr Trp Asp
 1               5                  10                  15
Val Pro Met Gly Gly Ile Trp Trp Asp Thr Ile Ala Gln Lys Ile Pro
            20                  25                  30
Asp Trp Ala Ser Ala Gly Ile Ser Ala Ile Trp Ile Pro Pro Ala Ser
        35                  40                  45
Lys Gly Met Ser Gly Gly Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Tyr
    50                  55                  60
Phe Asp Leu Gly Glu Tyr Tyr Gln Lys Gly Thr Val Glu Thr Arg Phe
65                  70                  75                  80
Gly Ser Lys Gln Glu Leu Ile Asn Met Ile Asn Thr Ala His Ala Tyr
                85                  90                  95
Gly Met Lys Val Ile Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly
            100                 105                 110
Asp Leu Glu Trp Asn Pro Phe Val Asn Asp Tyr Thr Trp Thr Asp Phe
        115                 120                 125
Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr Leu Asp Phe His
    130                 135                 140
Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe Gly Gly Tyr Pro
145                 150                 155                 160
Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp Leu Trp Ala Ser
                165                 170                 175
Gln Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly Ile Asp Ala Trp
            180                 185                 190
Arg Phe Asp Tyr Val Lys Gly Tyr Ala Pro Trp Val Val Lys Asp Trp
        195                 200                 205
Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn
    210                 215                 220
Val Asp Ala Val Leu Asn Trp Ala Tyr Ser Ser Gly Ala Lys Val Phe
225                 230                 235                 240
Asp Phe Ala Leu Tyr Tyr Lys Met Asp Glu Ala Phe Asp Asn Asn Asn
                245                 250                 255
Ile Pro Ala Leu Val Asp Ala Leu Arg Tyr Gly Gln Thr Val Val Ser
            260                 265                 270
Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Asp
        275                 280                 285
Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile Leu Thr Tyr Glu
    290                 295                 300
Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu Trp Leu Asn Lys
305                 310                 315                 320
Asp Lys Leu Lys Asn Leu Ile Trp Ile His Asp Asn Leu Ala Gly Gly
                325                 330                 335
Ser Thr Asp Ile Val Tyr Tyr Asp Asn Asp Glu Leu Ile Phe Val Arg
            340                 345                 350
Asn Gly Tyr Gly Ser Lys Pro Gly Leu Ile Thr Tyr Ile Asn Leu Ala
        355                 360                 365
Ser Ser Lys Ala Gly Arg Trp Val Tyr Val Pro Lys Phe Ala Gly Ser
    370                 375                 380
Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp Val Asp Lys Trp
385                 390                 395                 400
Val Asp Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro Ala His Asp Pro
                405                 410                 415
Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr Cys Gly Val Gly
            420                 425                 430
```

<210> 79
<211> 1386
<212> DNA
<213> Bacterial

<400> 79

```
atgaagcctg cgaaactcct cgtctttgtg ctcgtagtct ctatcctcgc ggggctctac      60
gcccagcccg cgggggcggc caagtacctg gagctcgaag agggcggcgt cataatgcag     120
gcgttctact gggacgtgcc ttcaggagga atatggtggg acacaatacg gcagaagata     180
ccggagtggt acgatgccgg aatctccgca atatggattc ccccggcgag caagggcatg     240
```

```
ggcggcgcct attcgatggg ctacgacccc tacgacttct ttgacctcgg tgagtacgac     300
cagaagggaa cggtagagac gcgctttggc tccaagcagg agctcgtgaa catgataaac     360
accgcccacg cctacggcat caaggtcatc gcagacatag taatcaacca ccgcgccgga     420
ggagaccttg agtggaaccc cttcgtcaat gactacacct ggacggactt ctcgaaggtc     480
gcttccggca agtacacggc caactacctc gacttccacc ccaacgaggt caagtgctgc     540
gacgagggca cctttggagg gttcccggac atagcccacg agaagagctg ggaccagtac     600
tggctctggg cgagcaacga gagctacgcc gcctacctca ggagcatcgg cgttgacgca     660
tggcgcttcg actacgtcaa gggctacgga gcgtgggtcg tcaaggactg gctggactgg     720
tggggaggct gggccgtcgg ggagtactgg gacacaaacg ttgatgcact gctcaactgg     780
gcctactcga gcgatgcaaa agtcttcgac ttcccgctct actacaagat ggacgcggcc     840
tttgacaaca agaacattcc cgcactcgtc gaggccctca agaacggggg cacagtcgtc     900
agccgcgacc cgtttaaggc cgtaaccttc gttgcaaacc acgacacgga cataatttgg     960
aacaagtacc cggcctacgc cttcatcctc acctacgagg gccagccgac gatattctac    1020
cgcgactacg aggagtggct caacaaggac aggctcaaga acctcatctg gatacacgac    1080
cacctcgccg gtggaagcac cgacatagtc tactacgata acgatgaact catcttcgtc    1140
aggaacggct acggggacaa gccggggctt ataacctaca tcaacctagg ctcgagcaag    1200
gccgggaggt gggtctacgt tccgaagttc gcgggagcgt gcatccacga gtacaccggc    1260
aacctcggcg gctgggtgga caagtgggtg gactcaagcg ggtgggtgta cctcgaggcc    1320
cctgcccacg acccggccaa cggctattac ggctactccg tctggagcta ctgcggggtg    1380
ggctga                                                               1386

<210> 80
<211> 461
<212> PRT
<213> Bacterial

<400> 80
Met Lys Pro Ala Lys Leu Leu Val Phe Val Leu Val Val Ser Ile Leu
 1               5                  10                  15
Ala Gly Leu Tyr Ala Gln Pro Ala Gly Ala Ala Lys Tyr Leu Glu Leu
            20                  25                  30
Glu Glu Gly Gly Val Ile Met Gln Ala Phe Tyr Trp Asp Val Pro Ser
        35                  40                  45
Gly Gly Ile Trp Trp Asp Thr Ile Arg Gln Lys Ile Pro Glu Trp Tyr
    50                  55                  60
Asp Ala Gly Ile Ser Ala Ile Trp Ile Pro Pro Ala Ser Lys Gly Met
65                  70                  75                  80
Gly Gly Ala Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Phe Phe Asp Leu
                85                  90                  95
Gly Glu Tyr Asp Gln Lys Gly Thr Val Glu Thr Arg Phe Gly Ser Lys
            100                 105                 110
Gln Glu Leu Val Asn Met Ile Asn Thr Ala His Ala Tyr Gly Ile Lys
        115                 120                 125
Val Ile Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly Asp Leu Glu
    130                 135                 140
Trp Asn Pro Phe Val Asn Asp Tyr Thr Trp Thr Asp Phe Ser Lys Val
145                 150                 155                 160
Ala Ser Gly Lys Tyr Thr Ala Asn Tyr Leu Asp Phe His Pro Asn Glu
                165                 170                 175
Val Lys Cys Cys Asp Glu Gly Thr Phe Gly Gly Phe Pro Asp Ile Ala
            180                 185                 190
His Glu Lys Ser Trp Asp Gln Tyr Trp Leu Trp Ala Ser Asn Glu Ser
        195                 200                 205
Tyr Ala Ala Tyr Leu Arg Ser Ile Gly Val Asp Ala Trp Arg Phe Asp
    210                 215                 220
Tyr Val Lys Gly Tyr Gly Ala Trp Val Val Lys Asp Trp Leu Asp Trp
225                 230                 235                 240
Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn Val Asp Ala
                245                 250                 255
Leu Leu Asn Trp Ala Tyr Ser Ser Asp Ala Lys Val Phe Asp Phe Pro
            260                 265                 270
Leu Tyr Tyr Lys Met Asp Ala Ala Phe Asp Asn Lys Asn Ile Pro Ala
        275                 280                 285
Leu Val Glu Ala Leu Lys Asn Gly Gly Thr Val Val Ser Arg Asp Pro
    290                 295                 300
Phe Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Asp Ile Ile Trp
305                 310                 315                 320
Asn Lys Tyr Pro Ala Tyr Ala Phe Ile Leu Thr Tyr Glu Gly Gln Pro
```

325 330 335
Thr Ile Phe Tyr Arg Asp Tyr Glu Glu Trp Leu Asn Lys Asp Arg Leu
340 345 350
Lys Asn Leu Ile Trp Ile His Asp His Leu Ala Gly Gly Ser Thr Asp
355 360 365
Ile Val Tyr Tyr Asp Asn Asp Glu Leu Ile Phe Val Arg Asn Gly Tyr
370 375 380
Gly Asp Lys Pro Gly Leu Ile Thr Tyr Ile Asn Leu Gly Ser Ser Lys
385 390 395 400
Ala Gly Arg Trp Val Tyr Val Pro Lys Phe Ala Gly Ala Cys Ile His
405 410 415
Glu Tyr Thr Gly Asn Leu Gly Gly Trp Val Asp Lys Trp Val Asp Ser
420 425 430
Ser Gly Trp Val Tyr Leu Glu Ala Pro Ala His Asp Pro Ala Asn Gly
435 440 445
Tyr Tyr Gly Tyr Ser Val Trp Ser Tyr Cys Gly Val Gly
450 455 460

<210> 81
<211> 1386
<212> DNA
<213> Bacterial

<400> 81
atgaagaagt ttgtcgccct gttcataacc atgtttttcg tagtgagcat ggcagtcgtt 60
gcacagccag ctagcgccgc aaagtattcc gagctcgaag aaggcggcgt tataatgcag 120
gccttctact gggacgtccc aggtggagga atctggtggg acaccatcag gagcaagata 180
ccggagtggt acgaggcggg aatatccgcc atttggattc cgccagccag caaggggatg 240
agcggcggtt actcgatggg ctacgatccc tacgatttct ttgacctcgg cgagtacaac 300
cagaagggaa ccatcgaaac gcgctttggc tctaaacagg agctcatcaa tatgataaac 360
acggcccatg cctacggcat aaaggtcata gcggacatcg tcataaacca ccgcgcaggc 420
ggagacctcg agtggaaccc gttcgttggg gactacacct ggacggactt ctcaaaggtg 480
gcctcgggca aatatactgc caactacctc gacttccacc ccaacgaggt caagtgctgt 540
gacgagggca catttggagg cttcccagac atagcccacg agaagagctg ggaccagcac 600
tggctctggg cgagcgatga gagctacgcc gcctacctaa ggagcatcgg cgttgatgcc 660
tggcgctttg actacgtgaa gggctacgga gcgtgggtcg tcaaggactg gctcaactgg 720
tggggcggct gggccgttgg cgagtactgg gacaccaacg ttgatgcact cctcaactgg 780
gcctactcga gcggcgccaa ggtcttcgac ttcccgctct actacaagat ggatgaggcc 840
tttgacaaca aaaacattcc agcgctcgtc tctgcccttc agaacggcca gactgttgtc 900
tcccgcgacc cgttcaaggc cgtaaccttt gtagcaaacc acgacaccga tataatctgg 960
aacaagtacc ttgcttatgc tttcatcctc acctacgaag gccagcccgt catattctac 1020
cgcgactacg aggagtggct caacaaggac aggttgaaca acctcatatg gatacacgac 1080
cacctcgcag gtggaagcac gagcatagtc tactacgaca gcgacgagat gatcttcgtg 1140
aggaacggct atggaagcaa gcctggcctt ataacttaca tcaacctcgg ctcgagcaag 1200
gttggaaggt gggtttatgt gccgaagttc gcgggcgcgt gcatccacga gtatactggt 1260
aacctcggag gctgggtaga caagtacgtc tactcaagcg gctgggtcta tctcgaagct 1320
ccagcttacg accctgccaa cgggcagtat ggctactccg tgtggagcta ttgcggtgtt 1380
gggtga 1386

<210> 82
<211> 461
<212> PRT
<213> Bacterial

<400> 82
Met Lys Lys Phe Val Ala Leu Phe Ile Thr Met Phe Phe Val Val Ser
1 5 10 15
Met Ala Val Val Ala Gln Pro Ala Ser Ala Ala Lys Tyr Ser Glu Leu
20 25 30
Glu Glu Gly Gly Val Ile Met Gln Ala Phe Tyr Trp Asp Val Pro Gly
35 40 45
Gly Gly Ile Trp Trp Asp Thr Ile Arg Ser Lys Ile Pro Glu Trp Tyr
50 55 60
Glu Ala Gly Ile Ser Ala Ile Trp Ile Pro Pro Ala Ser Lys Gly Met
65 70 75 80
Ser Gly Gly Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Phe Phe Asp Leu
85 90 95
Gly Glu Tyr Asn Gln Lys Gly Thr Ile Glu Thr Arg Phe Gly Ser Lys

```
            100                 105                 110
Gln Glu Leu Ile Asn Met Ile Asn Thr Ala His Ala Tyr Gly Ile Lys
        115                 120                 125
Val Ile Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly Asp Leu Glu
    130                 135                 140
Trp Asn Pro Phe Val Gly Asp Tyr Thr Trp Thr Asp Phe Ser Lys Val
145                 150                 155                 160
Ala Ser Gly Lys Tyr Thr Ala Asn Tyr Leu Asp Phe His Pro Asn Glu
                165                 170                 175
Val Lys Cys Cys Asp Glu Gly Thr Phe Gly Gly Phe Pro Asp Ile Ala
            180                 185                 190
His Glu Lys Ser Trp Asp Gln His Trp Leu Trp Ala Ser Asp Glu Ser
        195                 200                 205
Tyr Ala Ala Tyr Leu Arg Ser Ile Gly Val Asp Ala Trp Arg Phe Asp
    210                 215                 220
Tyr Val Lys Gly Tyr Gly Ala Trp Val Val Lys Asp Trp Leu Asn Trp
225                 230                 235                 240
Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn Val Asp Ala
                245                 250                 255
Leu Leu Asn Trp Ala Tyr Ser Ser Gly Ala Lys Val Phe Asp Phe Pro
            260                 265                 270
Leu Tyr Tyr Lys Met Asp Glu Ala Phe Asp Asn Lys Asn Ile Pro Ala
        275                 280                 285
Leu Val Ser Ala Leu Gln Asn Gly Gln Thr Val Val Ser Arg Asp Pro
    290                 295                 300
Phe Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Asp Ile Ile Trp
305                 310                 315                 320
Asn Lys Tyr Leu Ala Tyr Ala Phe Ile Leu Thr Tyr Glu Gly Gln Pro
                325                 330                 335
Val Ile Phe Tyr Arg Asp Tyr Glu Glu Trp Leu Asn Lys Asp Arg Leu
            340                 345                 350
Asn Asn Leu Ile Trp Ile His Asp His Leu Ala Gly Gly Ser Thr Ser
        355                 360                 365
Ile Val Tyr Tyr Asp Ser Asp Glu Met Ile Phe Val Arg Asn Gly Tyr
    370                 375                 380
Gly Ser Lys Pro Gly Leu Ile Thr Tyr Ile Asn Leu Gly Ser Ser Lys
385                 390                 395                 400
Val Gly Arg Trp Val Tyr Val Pro Lys Phe Ala Gly Ala Cys Ile His
                405                 410                 415
Glu Tyr Thr Gly Asn Leu Gly Gly Trp Val Asp Lys Tyr Val Tyr Ser
            420                 425                 430
Ser Gly Trp Val Tyr Leu Glu Ala Pro Ala Tyr Asp Pro Ala Asn Gly
        435                 440                 445
Gln Tyr Gly Tyr Ser Val Trp Ser Tyr Cys Gly Val Gly
    450                 455                 460
```

<210> 83
<211> 1299
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 83
atggctctgg aagacggcgg gctcataatg caggccttct actgggatgt tcctggagga 60
ggaatctggt gggacacaat agctcaaaag atacccgaat gggcaagtgc aggaatctca 120
gcgatatgga ttccaccagc gagtaagggc atgagcggtg gttattccat gggctacgat 180
ccctacgatt tctttgacct cggcgagtac tatcagaagg ggacagttga gacgcgcttc 240
ggctcaaagg aagaactggt gaacatgata aacaccgcac actcctacgg cataaaggtg 300
atagcagaca tagtcataaa ccaccgcgcc ggtggagacc ttgagtggaa ccccttcgtg 360
aacgactata cctggacaga cttctcaaaa gtcgcctccg gtaaatatac ggccaactac 420
cttgacttcc acccaaacga gcttcactgt tgtgatgaag gtacctttgg aggataccct 480
gatatatgtc acgacaaaag ctgggaccag tactggctct gggcgagcag cgaaagctac 540
gctgcctacc tcaggagcat aggggttgac gcctggcgtt tcgactacgt caagggctac 600
ggagcatggg ttgttaacga ctggctcagc tggtggggag gctgggccgt tggagagtac 660
tgggacacga acgttgatgc actcctcaac tgggcataca gcagcggcgc caaggtcttt 720
gacttcccgc tctactacaa gatggacgaa gccttcgaca acaccaacat cccggcatta 780

gtggatgcac tcagatacgg ccagacagtg gtcagccgcg atcccttcaa ggcggtaact 840
ttcgttgcca accacgatac agatataatc tggaacaagt atccggctta tgcattcatc 900
cttacctatg agggacagcc tgttatattc taccgcgact acgaggagtg gctcaacaag 960
gataagctta acaacctcat ctggatacac gatcaccttg ctggagggag tactgacatt 1020
gtttactacg acagcgacga gcttatcttt gtgagaaacg gctatggcac caaaccagga 1080
ctgataacct atatcaacct cggctcaagc aaagttggaa ggtgggtcta cgttccaaag 1140
ttcgccggtt catgcatcca cgagtacacc ggcaacctcg gcggttggat agacaagtac 1200
gtctcctcca gcggctgggt ctatcttgag gccccagccc acgacccggc gaacggctac 1260
tacggctact ccgtatggag ctactgcggg gttgggtga 1299

<210> 84
<211> 432
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 84
Met Ala Leu Glu Asp Gly Gly Leu Ile Met Gln Ala Phe Tyr Trp Asp
1 5 10 15
Val Pro Gly Gly Gly Ile Trp Trp Asp Thr Ile Ala Gln Lys Ile Pro
20 25 30
Glu Trp Ala Ser Ala Gly Ile Ser Ala Ile Trp Ile Pro Pro Ala Ser
35 40 45
Lys Gly Met Ser Gly Gly Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Phe
50 55 60
Phe Asp Leu Gly Glu Tyr Tyr Gln Lys Gly Thr Val Glu Thr Arg Phe
65 70 75 80
Gly Ser Lys Glu Glu Leu Val Asn Met Ile Asn Thr Ala His Ser Tyr
85 90 95
Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly
100 105 110
Asp Leu Glu Trp Asn Pro Phe Val Asn Asp Tyr Thr Trp Thr Asp Phe
115 120 125
Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr Leu Asp Phe His
130 135 140
Pro Asn Glu Leu His Cys Cys Asp Glu Gly Thr Phe Gly Gly Tyr Pro
145 150 155 160
Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp Leu Trp Ala Ser
165 170 175
Ser Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly Val Asp Ala Trp
180 185 190
Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp Val Val Asn Asp Trp
195 200 205
Leu Ser Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn
210 215 220
Val Asp Ala Leu Leu Asn Trp Ala Tyr Ser Ser Gly Ala Lys Val Phe
225 230 235 240
Asp Phe Pro Leu Tyr Tyr Lys Met Asp Glu Ala Phe Asp Asn Thr Asn
245 250 255
Ile Pro Ala Leu Val Asp Ala Leu Arg Tyr Gly Gln Thr Val Val Ser
260 265 270
Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Asp
275 280 285
Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile Leu Thr Tyr Glu
290 295 300
Gly Gln Pro Val Ile Phe Tyr Arg Asp Tyr Glu Glu Trp Leu Asn Lys
305 310 315 320
Asp Lys Leu Asn Asn Leu Ile Trp Ile His Asp His Leu Ala Gly Gly
325 330 335
Ser Thr Asp Ile Val Tyr Tyr Asp Ser Asp Glu Leu Ile Phe Val Arg
340 345 350
Asn Gly Tyr Gly Thr Lys Pro Gly Leu Ile Thr Tyr Ile Asn Leu Gly
355 360 365
Ser Ser Lys Val Gly Arg Trp Val Tyr Val Pro Lys Phe Ala Gly Ser
370 375 380
Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp Ile Asp Lys Tyr

385 390 395 400
Val Ser Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro Ala His Asp Pro
405 410 415
Ala Asn Gly Tyr Tyr Gly Tyr Ser Val Trp Ser Tyr Cys Gly Val Gly
420 425 430

<210> 85
<211> 1299
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 85
atggctctgg aagagggcgg gcttataatg caggcattct attgggacgt cccaggtgga 60
ggaatctggt gggacaccat agcccagaag atacccgaat gggcaagtgc aggaatctca 120
gcgatatgga ttccaccagc gagtaaggga atgagcggtg gttattccat gggctacgat 180
ccctacgatt tctttgacct cggcgagtac tatcagaagg ggacagttga gacgcgcttc 240
ggctcaaagg aagaactggt gaacatgata aacaccgcac actcctacgg cataaaggtg 300
atagcggaca tagtcataaa ccaccgcgcc ggtggaggcc tcgagtggaa ccccttcgtg 360
aacgactata cctggacaga cttctcaaaa gtcgcctccg gtaaatatac agccaactac 420
cttgacttcc acccaaacga gcttcactgt tgtgatgaag gtacctttgg aggataccct 480
gatatatgtc acgacaaaag ctgggaccag tactggctct gggcgagcag cgaaagctac 540
gctgcctacc tcaggagcat aggggttgac gcctggtgtt tcgactacgt caagggctac 600
ggagcctggg ttgttaacga ctggctcagc tggtggggag gctgggccgt tggagagtac 660
tgggacacta acgttgatgc actcctcaac tgggcataca acagcggcgc caaggtcttt 720
gacttcccgc tctactacaa gatggacgaa gccttcgaca ataccaacat ccccgctttg 780
gtttacgccc tcaagaatgg cgggacagtg gtcagccgcg acccattcaa ggcggtaact 840
ttcgttgcca accacgatac agatataatc tggaacaagt atccggctta tgcattcatc 900
cttacctatg agggacagcc tgttatattc taccgcgact acgaggagtg gctcaacaag 960
gataagctta acaacctcat ctggatacac gatcaccttg ctggagggag tactgacatt 1020
gtttactacg acagcgacga gcttatcttt gtgagaaacg gctatggcac caaaccagga 1080
ctgataacct atatcaacct cggctcaagc aaagctggaa ggtgggtcta cgttccaaag 1140
ttcgccggtt catgcatcca cgagtacacc ggcagcctcg gcggttggat agacaagtac 1200
gtctcctcca gcggctgggt ctaccttgag gccccggccc acgacccggc caatggccag 1260
tatggctact ccgtctggag ctattgcggg gttgggtga 1299

<210> 86
<211> 432
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 86
Met Ala Leu Glu Glu Gly Gly Leu Ile Met Gln Ala Phe Tyr Trp Asp
1 5 10 15
Val Pro Gly Gly Gly Ile Trp Trp Asp Thr Ile Ala Gln Lys Ile Pro
20 25 30
Glu Trp Ala Ser Ala Gly Ile Ser Ala Ile Trp Ile Pro Pro Ala Ser
35 40 45
Lys Gly Met Ser Gly Gly Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Phe
50 55 60
Phe Asp Leu Gly Glu Tyr Tyr Gln Lys Gly Thr Val Glu Thr Arg Phe
65 70 75 80
Gly Ser Lys Glu Glu Leu Val Asn Met Ile Asn Thr Ala His Ser Tyr
85 90 95
Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly
100 105 110
Gly Leu Glu Trp Asn Pro Phe Val Asn Asp Tyr Thr Trp Thr Asp Phe
115 120 125
Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr Leu Asp Phe His
130 135 140
Pro Asn Glu Leu His Cys Cys Asp Glu Gly Thr Phe Gly Gly Tyr Pro
145 150 155 160
Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp Leu Trp Ala Ser

```
                165                 170                 175
Ser Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly Val Asp Ala Trp
            180                 185                 190
Cys Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp Val Val Asn Asp Trp
        195                 200                 205
Leu Ser Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn
    210                 215                 220
Val Asp Ala Leu Leu Asn Trp Ala Tyr Asn Ser Gly Ala Lys Val Phe
225                 230                 235                 240
Asp Phe Pro Leu Tyr Tyr Lys Met Asp Glu Ala Phe Asp Asn Thr Asn
                245                 250                 255
Ile Pro Ala Leu Val Tyr Ala Leu Lys Asn Gly Gly Thr Val Val Ser
            260                 265                 270
Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Asp
        275                 280                 285
Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile Leu Thr Tyr Glu
    290                 295                 300
Gly Gln Pro Val Ile Phe Tyr Arg Asp Tyr Glu Glu Trp Leu Asn Lys
305                 310                 315                 320
Asp Lys Leu Asn Asn Leu Ile Trp Ile His Asp His Leu Ala Gly Gly
                325                 330                 335
Ser Thr Asp Ile Val Tyr Tyr Asp Ser Asp Glu Leu Ile Phe Val Arg
            340                 345                 350
Asn Gly Tyr Gly Thr Lys Pro Gly Leu Ile Thr Tyr Ile Asn Leu Gly
        355                 360                 365
Ser Ser Lys Ala Gly Arg Trp Val Tyr Val Pro Lys Phe Ala Gly Ser
    370                 375                 380
Cys Ile His Glu Tyr Thr Gly Ser Leu Gly Gly Trp Ile Asp Lys Tyr
385                 390                 395                 400
Val Ser Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro Ala His Asp Pro
                405                 410                 415
Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr Cys Gly Val Gly
            420                 425                 430
```

<210> 87
<211> 1419
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 87

```
atgttcctgc tcgcgttttt gctcactgcc tcgctgttct gcccaacagg acagcccgcc      60
aaggctgccg caccgtttaa cggcaccatg atgcagtatt ttgaatggta cttgccggat     120
gatggcacgt tatggaccaa agtggccaat gaagccaaca acttatccag ccttggcatc     180
accgctcttt ggctgccgcc cgcttacaaa ggaacaagcc gcagcgacgt agggtacgga     240
gtatacgact tgtatgacct cggcgaattc aatcaaaaag ggaccgtccg cacaaaatac     300
ggaacaaaag ctcaatatct tcaagccatt caagccgccc acgccgctgg aatgcaagtg     360
tacgccgatg tcgtgttcga ccataaaggc ggcgctgacg gcacggaatg ggtggacgcc     420
gtcgaagtca atccgtccga ccgcaaccaa gaaatctcgg gcacctatca aatccaagca     480
tggacgaaat ttgattttcc cgggcggggc aacacctact ccagctttaa gtggcgctgg     540
taccattttg acggcgttga ttgggacgaa agccgaaaat tgagccgcat ttacaaattc     600
cgcggcatcg gcaaagcgtg ggattgggaa gtagacacgg aaaacggaaa ctatgactac     660
ttaatgtatg ccgaccttga tatggatcat cccgaagtcg tgaccgagct gaaaaactgg     720
gggaaatggt atgtcaacac aacgaacatt gatgggttcc ggcttgatgc cgtcaagcat     780
attaagttca gttttttcc tgattggttg tcgtatgtgc gttctcagac tggcaagccg     840
ctatttaccg tcggggaata ttggagctat gacatcaaca agttgcacaa ttacattacg     900
aaaacagacg gaacgatgtc tttgtttgat gccccgttac acaacaaatt ttataccgct     960
tccaaatcag ggggcgcatt tgatatgcgc acgttaatga ccaatactct catgaaagat    1020
caaccgacat tggccgtcac cttcgttgat aatcatgaca ccgaacccgg ccaagcgctg    1080
cagtcatggg tcgacccatg gttcaaaccg ttggcttacg cctttattct aactcggcag    1140
gaaggatacc cgtgcgtctt ttatggtgac tattatggca ttccacaata taacattcct    1200
tcgctgaaaa gcaaaatcga tccgctcctc atcgcgcgca gggattatgc ttacggaacg    1260
caacatgatt atcttgatca ctccgacatc atcgggtgga caagggaagg ggtcactgaa    1320
aaaccaggat ccgggctggc cgcactgatc accgatgggc cgggaggaag caaatggatg    1380
tactgttggc aaacaacacg ctggaaaagt gttctatga                           1419
```

<210> 88
<211> 472
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 88
Met Phe Leu Leu Ala Phe Leu Leu Thr Ala Ser Leu Phe Cys Pro Thr
 1   5   10   15
Gly Gln Pro Ala Lys Ala Ala Ala Pro Phe Asn Gly Thr Met Met Gln
            20   25   30
Tyr Phe Glu Trp Tyr Leu Pro Asp Asp Gly Thr Leu Trp Thr Lys Val
        35   40   45
Ala Asn Glu Ala Asn Asn Leu Ser Ser Leu Gly Ile Thr Ala Leu Trp
    50   55   60
Leu Pro Pro Ala Tyr Lys Gly Thr Ser Arg Ser Asp Val Gly Tyr Gly
65   70   75   80
Val Tyr Asp Leu Tyr Asp Leu Gly Glu Phe Asn Gln Lys Gly Thr Val
                85   90   95
Arg Thr Lys Tyr Gly Thr Lys Ala Gln Tyr Leu Gln Ala Ile Gln Ala
            100   105   110
Ala His Ala Ala Gly Met Gln Val Tyr Ala Asp Val Val Phe Asp His
        115   120   125
Lys Gly Gly Ala Asp Gly Thr Glu Trp Val Asp Ala Val Glu Val Asn
    130   135   140
Pro Ser Asp Arg Asn Gln Glu Ile Ser Gly Thr Tyr Gln Ile Gln Ala
145   150   155   160
Trp Thr Lys Phe Asp Phe Pro Gly Arg Gly Asn Thr Tyr Ser Ser Phe
                165   170   175
Lys Trp Arg Trp Tyr His Phe Asp Gly Val Asp Trp Asp Glu Ser Arg
            180   185   190
Lys Leu Ser Arg Ile Tyr Lys Phe Arg Gly Ile Gly Lys Ala Trp Asp
        195   200   205
Trp Glu Val Asp Thr Glu Asn Gly Asn Tyr Asp Tyr Leu Met Tyr Ala
    210   215   220
Asp Leu Asp Met Asp His Pro Glu Val Val Thr Glu Leu Lys Asn Trp
225   230   235   240
Gly Lys Trp Tyr Val Asn Thr Thr Asn Ile Asp Gly Phe Arg Leu Asp
                245   250   255
Ala Val Lys His Ile Lys Phe Ser Phe Phe Pro Asp Trp Leu Ser Tyr
            260   265   270
Val Arg Ser Gln Thr Gly Lys Pro Leu Phe Thr Val Gly Glu Tyr Trp
        275   280   285
Ser Tyr Asp Ile Asn Lys Leu His Asn Tyr Ile Thr Lys Thr Asp Gly
    290   295   300
Thr Met Ser Leu Phe Asp Ala Pro Leu His Asn Lys Phe Tyr Thr Ala
305   310   315   320
Ser Lys Ser Gly Gly Ala Phe Asp Met Arg Thr Leu Met Thr Asn Thr
                325   330   335
Leu Met Lys Asp Gln Pro Thr Leu Ala Val Thr Phe Val Asp Asn His
            340   345   350
Asp Thr Glu Pro Gly Gln Ala Leu Gln Ser Trp Val Asp Pro Trp Phe
        355   360   365
Lys Pro Leu Ala Tyr Ala Phe Ile Leu Thr Arg Gln Glu Gly Tyr Pro
    370   375   380
Cys Val Phe Tyr Gly Asp Tyr Tyr Gly Ile Pro Gln Tyr Asn Ile Pro
385   390   395   400
Ser Leu Lys Ser Lys Ile Asp Pro Leu Leu Ile Ala Arg Arg Asp Tyr
                405   410   415
Ala Tyr Gly Thr Gln His Asp Tyr Leu Asp His Ser Asp Ile Ile Gly
            420   425   430
Trp Thr Arg Glu Gly Val Thr Glu Lys Pro Gly Ser Gly Leu Ala Ala
        435   440   445
Leu Ile Thr Asp Gly Pro Gly Gly Ser Lys Trp Met Tyr Cys Trp Gln
    450   455   460
Thr Thr Arg Trp Lys Ser Val Leu

465 470

<210> 89
<211> 3301
<212> DNA
<213> Bacterial

<400> 89
atgaaagaag cggttgtgta tcaaattttc ccggatcggt tctttaatgg caacccttca 60
aatgataaca gcaagcagca ggcacgcggg gcgcagccga ttgagcatcg cgattggtcg 120
gatttgcccg ataatccgcg cctgaaaggg acgagcggct acgatggcga cggtgaatgg 180
tcgaatgact ttttcggcgg agacatcgcc ggaattgaac aaaagttgga ttatttgcag 240
tcgcttggag tgaacacgat ttacttaaat ccgatcgcca atgcgccatc gaaccataaa 300
tatgatgcga gcaattacaa agaattggat ccgatgttcg gttccccgga agaattccaa 360
tcgtttgtgc aggcgcttgc gaaccggggg atgcatctca tcttagacgg ggtgttcaac 420
cacgtatccg acgattcgat ttactttgac cgctaccacc gctatccgac cgtcggtgcg 480
tatgaatatt gggaagcggt ttacgatttg atgaatgaaa aaggattgag cgaggaagaa 540
gcgcggaaac aagtggaaga gaagttcaaa caagagggac agacgttcag cccgtatggg 600
tttcatcttt ggttcaatat tgaaaacaaa aaagtcaatg gccattatca ataccaatca 660
tggtggggct atgacagtct gccggagttt aagtcggtga cgggggaaaa agtgccgcat 720
ccgagtgaat tgaacaacga tgcgctcgcg aattacattt tccgtgaatc ggattcggtg 780
gcgaaaagct ggattgccct cggcgcctcc ggctggcggt tggatgtggc caatgaggtg 840
gatccggcgt tttggcgcga gtttcgccaa gaattgcttc aagggtcgta cggccgcggt 900
ccgacgttaa aagaggggga gcagccgctc attttagggg aaatttggga tgacgcatcg 960
aaatattttc taggcgacca gtacgattcc gtgatgaact accggttccg cggggcggtg 1020
cttgactttt tgaaaaacgg aaatgcagaa gaggcggaca agcggctgac ggccataagg 1080
gaagactacc caagtgaagc gttttatgcg ctgatgaact taatcggttc gcatgacacg 1140
gcgcgggcgg tctttctgct tgggaacgga acggattcat ccgagcgggc ggagcttgat 1200
ccgaattata atgaggaact tgggaaaaag cggctcaagc tggcggtgat tttgcagatg 1260
ggatacccgg gagcgccgac gatttattac ggcgatgaag cgggagtaac aggctcaaaa 1320
gacccagaca accgccgcac gtatccgtgg ggcaaagaag atcaaaatct gttgtcccat 1380
tatcagaaag tggggcacat tcgccagcac catcaatcgt tgttggccca tggcgacatc 1440
aagacggtgt atgcgcaagg ggatgtatac gtatttgccc gccaatacgg gcgtgaagcg 1500
gcgctcattg ccatcaaccg cggcaatgag gacaagacgg tggcgcttga cgtcgcttcg 1560
ttgcttccga acggcaccgt gcttacggat gagttgcatg atggcgggga agctacggtc 1620
gctggcggaa cgttgacggt cacgattccg gccctggatg gacggatgat gtttgggacg 1680
gtgacggcgg aaatgccggc agcagtcagc aatttgcagg cgagcgcttc ggatggctgc 1740
gtgacgttaa cgtgggaagg aaatgcatcg agataccgaa tttacgagtc cacgttaaaa 1800
ggtgccggtt atacgatggt gcaagagacg gaaacaactt cggccacgat cggttcgttg 1860
acgaacggaa cagcctatta ctttgccgtt gcggcggtcg atgaaaacgg gaatgaatca 1920
ccgaaggtcg aaacgaatcg cgtcgttcct cattacccgc tgacgagcga caatgtccag 1980
ttcgtgacaa cgttaagcga tgccacactg gatttgtcaa agccgcagca agtggatgtc 2040
catgtcaaca tcgacaatgt gacaagcaaa ggagcagctg atgggttgca agcggtgttg 2100
caagtgaaag gcccgcatga cgaaacatgg aaagaataca gagcggctta ccaaggacaa 2160
gacggcgacg ccaacgtgtt ccgagctgcc ttcactccgc tcgccgcagg gacgtatacg 2220
tatcgttatg cgctgacgac caaccttggc gaggagtgga tgtatacaga agagaagcaa 2280
gtgacgtttg cggcagacaa cagcgaccaa atagcgccag cagacgccat cgagctgcgg 2340
cagcctgcgg ttgaatcggg acaagtgaat ttatcatgga cgtttgttgg gaaaaaagat 2400
ggggatgctt atttgttagc catcgagcgc aacggtgata tcgtgcatac aaccacttcg 2460
atcggcgatt catttacaga ctacgatgtc gaaaacggca ccgagtacac gtatgttgtc 2520
aagttgtatg accgcgccgg caatgttgtg gcgtcaaaca cggtcaaggt gacgccggac 2580
attgtgatgg tgaaagtgat ttttaaagtg agagcgccgg attacacacc gttggatgcc 2640
cgaattacga ttccgaacag cttgaacggc tggaacacag gggcctggga gatgtcgcgc 2700
aacggtgcgg tgacgcccga ttggcaattt accgtcgagg tgcaggaagg ggaaacgatc 2760
acctataagt atgtgaaagg cggatcgtgg gatcaagagg ggttggccga ccatacgcgt 2820
gaggacgaca acgatgatga cgtgagctac tacggctatg ggacgattgg caccgacttg 2880
aaagtgacgg tccacaatga aggaaacaat acgatgattg tgcaagaccg catttgcgc 2940
tggatcgata tgccggtcgt catcgaagag gtgcaaaaac aaggaagtca agtgacgatc 3000
aagggcaatg ccattaaaaa cggtgttttg acgatcaatg gcgagcgggt gccgattgat 3060
ggccggatgg cattctcgta cacgtttgcg ccggccagcc atcaaaaaga agtgttgatc 3120
catatcgaac catcggccga aagcaaaaca gccattttca acaacgacgg cggagcgatt 3180
gcgaaaaaca caaaagatta cgtgctgaat ttagaaacga agcaattcaa aaagcttctc 3240
gagagtactt ctagagcggc cgcgggccca tcgattttcc acccgggtgg ggtaccaggt 3300
a 3301

<210> 90
<211> 1100
<212> PRT

<213> Bacterial

<400> 90

```
Met Lys Glu Ala Val Val Tyr Gln Ile Phe Pro Asp Arg Phe Phe Asn
 1               5                  10                  15
Gly Asn Pro Ser Asn Asp Asn Ser Lys Gln Gln Ala Arg Gly Ala Gln
            20                  25                  30
Pro Ile Glu His Arg Asp Trp Ser Asp Leu Pro Asp Asn Pro Arg Leu
        35                  40                  45
Lys Gly Thr Ser Gly Tyr Asp Gly Asp Gly Glu Trp Ser Asn Asp Phe
    50                  55                  60
Phe Gly Gly Asp Ile Ala Gly Ile Glu Gln Lys Leu Asp Tyr Leu Gln
65                  70                  75                  80
Ser Leu Gly Val Asn Thr Ile Tyr Leu Asn Pro Ile Ala Asn Ala Pro
                85                  90                  95
Ser Asn His Lys Tyr Asp Ala Ser Asn Tyr Lys Glu Leu Asp Pro Met
            100                 105                 110
Phe Gly Ser Pro Glu Glu Phe Gln Ser Phe Val Gln Ala Leu Ala Asn
        115                 120                 125
Arg Gly Met His Leu Ile Leu Asp Gly Val Phe Asn His Val Ser Asp
    130                 135                 140
Asp Ser Ile Tyr Phe Asp Arg Tyr His Arg Tyr Pro Thr Val Gly Ala
145                 150                 155                 160
Tyr Glu Tyr Trp Glu Ala Val Tyr Asp Leu Met Asn Glu Lys Gly Leu
                165                 170                 175
Ser Glu Glu Glu Ala Arg Lys Gln Val Glu Glu Lys Phe Lys Gln Glu
            180                 185                 190
Gly Gln Thr Phe Ser Pro Tyr Gly Phe His Leu Trp Phe Asn Ile Glu
        195                 200                 205
Asn Lys Lys Val Asn Gly His Tyr Gln Tyr Gln Ser Trp Trp Gly Tyr
    210                 215                 220
Asp Ser Leu Pro Glu Phe Lys Ser Val Thr Gly Glu Lys Val Pro His
225                 230                 235                 240
Pro Ser Glu Leu Asn Asn Asp Ala Leu Ala Asn Tyr Ile Phe Arg Glu
                245                 250                 255
Ser Asp Ser Val Ala Lys Ser Trp Ile Ala Leu Gly Ala Ser Gly Trp
            260                 265                 270
Arg Leu Asp Val Ala Asn Glu Val Asp Pro Ala Phe Trp Arg Glu Phe
        275                 280                 285
Arg Gln Glu Leu Leu Gln Gly Ser Tyr Gly Arg Gly Pro Thr Leu Lys
    290                 295                 300
Glu Gly Glu Gln Pro Leu Ile Leu Gly Glu Ile Trp Asp Asp Ala Ser
305                 310                 315                 320
Lys Tyr Phe Leu Gly Asp Gln Tyr Asp Ser Val Met Asn Tyr Arg Phe
                325                 330                 335
Arg Gly Ala Val Leu Asp Phe Leu Lys Asn Gly Asn Ala Glu Glu Ala
            340                 345                 350
Asp Lys Arg Leu Thr Ala Ile Arg Glu Asp Tyr Pro Ser Glu Ala Phe
        355                 360                 365
Tyr Ala Leu Met Asn Leu Ile Gly Ser His Asp Thr Ala Arg Ala Val
    370                 375                 380
Phe Leu Leu Gly Asn Gly Thr Asp Ser Ser Glu Arg Ala Glu Leu Asp
385                 390                 395                 400
Pro Asn Tyr Asn Glu Glu Leu Gly Lys Lys Arg Leu Lys Leu Ala Val
                405                 410                 415
Ile Leu Gln Met Gly Tyr Pro Gly Ala Pro Thr Ile Tyr Tyr Gly Asp
            420                 425                 430
Glu Ala Gly Val Thr Gly Ser Lys Asp Pro Asp Asn Arg Arg Thr Tyr
        435                 440                 445
Pro Trp Gly Lys Glu Asp Gln Asn Leu Leu Ser His Tyr Gln Lys Val
    450                 455                 460
Gly His Ile Arg Gln His His Gln Ser Leu Leu Ala His Gly Asp Ile
465                 470                 475                 480
Lys Thr Val Tyr Ala Gln Gly Asp Val Tyr Val Phe Ala Arg Gln Tyr
                485                 490                 495
Gly Arg Glu Ala Ala Leu Ile Ala Ile Asn Arg Gly Asn Glu Asp Lys
            500                 505                 510
Thr Val Ala Leu Asp Val Ala Ser Leu Leu Pro Asn Gly Thr Val Leu
```

```
        515                 520                 525
Thr Asp Glu Leu His Asp Gly Gly Glu Ala Thr Val Ala Gly Gly Thr
    530                 535                 540
Leu Thr Val Thr Ile Pro Ala Leu Asp Gly Arg Met Met Phe Gly Thr
545                 550                 555                 560
Val Thr Ala Glu Met Pro Ala Ala Val Ser Asn Leu Gln Ala Ser Ala
                565                 570                 575
Ser Asp Gly Cys Val Thr Leu Thr Trp Glu Gly Asn Ala Ser Arg Tyr
            580                 585                 590
Arg Ile Tyr Glu Ser Thr Leu Lys Gly Ala Gly Tyr Thr Met Val Gln
        595                 600                 605
Glu Thr Glu Thr Thr Ser Ala Thr Ile Gly Ser Leu Thr Asn Gly Thr
    610                 615                 620
Ala Tyr Tyr Phe Ala Val Ala Ala Val Asp Glu Asn Gly Asn Glu Ser
625                 630                 635                 640
Pro Lys Val Glu Thr Asn Arg Val Val Pro His Tyr Pro Leu Thr Ser
                645                 650                 655
Asp Asn Val Gln Phe Val Thr Thr Leu Ser Asp Ala Thr Leu Asp Leu
            660                 665                 670
Ser Lys Pro Gln Gln Val Asp Val His Val Asn Ile Asp Asn Val Thr
        675                 680                 685
Ser Lys Gly Ala Ala Asp Gly Leu Gln Ala Val Leu Gln Val Lys Gly
    690                 695                 700
Pro His Asp Glu Thr Trp Lys Glu Tyr Arg Ala Ala Tyr Gln Gly Gln
705                 710                 715                 720
Asp Gly Asp Ala Asn Val Phe Arg Ala Ala Phe Thr Pro Leu Ala Ala
                725                 730                 735
Gly Thr Tyr Thr Tyr Arg Tyr Ala Leu Thr Thr Asn Leu Gly Glu Glu
            740                 745                 750
Trp Met Tyr Thr Glu Glu Lys Gln Val Thr Phe Ala Ala Asp Asn Ser
        755                 760                 765
Asp Gln Ile Ala Pro Ala Asp Ala Ile Glu Leu Arg Gln Pro Ala Val
    770                 775                 780
Glu Ser Gly Gln Val Asn Leu Ser Trp Thr Phe Val Gly Lys Lys Asp
785                 790                 795                 800
Gly Asp Ala Tyr Leu Leu Ala Ile Glu Arg Asn Gly Asp Ile Val His
                805                 810                 815
Thr Thr Thr Ser Ile Gly Asp Ser Phe Thr Asp Tyr Asp Val Glu Asn
            820                 825                 830
Gly Thr Glu Tyr Thr Tyr Val Val Lys Leu Tyr Asp Arg Ala Gly Asn
        835                 840                 845
Val Val Ala Ser Asn Thr Val Lys Val Thr Pro Asp Ile Val Met Val
    850                 855                 860
Lys Val Ile Phe Lys Val Arg Ala Pro Asp Tyr Thr Pro Leu Asp Ala
865                 870                 875                 880
Arg Ile Thr Ile Pro Asn Ser Leu Asn Gly Trp Asn Thr Gly Ala Trp
                885                 890                 895
Glu Met Ser Arg Asn Gly Ala Val Thr Pro Asp Trp Gln Phe Thr Val
            900                 905                 910
Glu Val Gln Glu Gly Glu Thr Ile Thr Tyr Lys Tyr Val Lys Gly Gly
        915                 920                 925
Ser Trp Asp Gln Glu Gly Leu Ala Asp His Thr Arg Glu Asp Asp Asn
    930                 935                 940
Asp Asp Asp Val Ser Tyr Tyr Gly Tyr Gly Thr Ile Gly Thr Asp Leu
945                 950                 955                 960
Lys Val Thr Val His Asn Glu Gly Asn Asn Thr Met Ile Val Gln Asp
                965                 970                 975
Arg Ile Leu Arg Trp Ile Asp Met Pro Val Val Ile Glu Glu Val Gln
            980                 985                 990
Lys Gln Gly Ser Gln Val Thr Ile Lys Gly Asn Ala Ile Lys Asn Gly
        995                 1000                1005
Val Leu Thr Ile Asn Gly Glu Arg Val Pro Ile Asp Gly Arg Met Ala
    1010                1015                1020
Phe Ser Tyr Thr Phe Ala Pro Ala Ser His Gln Lys Glu Val Leu Ile
1025                1030                1035                1040
His Ile Glu Pro Ser Ala Glu Ser Lys Thr Ala Ile Phe Asn Asn Asp
                1045                1050                1055
Gly Gly Ala Ile Ala Lys Asn Thr Lys Asp Tyr Val Leu Asn Leu Glu
```

1060 1065 1070
Thr Lys Gln Phe Lys Lys Leu Leu Glu Ser Thr Ser Arg Ala Ala Ala
1075 1080 1085
Gly Pro Ser Ile Phe His Pro Gly Gly Val Pro Gly
1090 1095 1100

<210> 91
<211> 1650
<212> DNA
<213> Bacterial

<400> 91
gtgctaacgt ttcaccgcat cattcgaaaa ggatggatgt tcctgctcgc gtttttgctc 60
actgcctcgc tgttctgccc aacaggacag cccgccaagg ctgccgcacc gtttaacggc 120
accatgatgc agtattttga atggtacttg ccggatgatg gcacgttatg gaccaaagtg 180
gccaatgaag ccaacaactt atccagcctt ggcatcaccg ctctttggct gccgcccgct 240
tataaaggaa caagccgcag cgacgtaggg tacggagtat acgacttgta tgacctcggc 300
gaattcaatc aaaaagggac cgtccgcaca aaatacggaa caaaagctca atatcttcaa 360
gccattcaag ccgcccacgc cgctggaatg caagtgtacg ccgatgtcgt gttcgaccat 420
aaaggcggcg ccgacggcac ggaatgggtg gacgccgtcg aagtcaatcc gtccgaccgc 480
aaccaagaaa tctcgggcac ctatcaaatc caagcatgga cgaaatttga ttttcccggg 540
cggggcaaca cctactccag ctttaagtgg cgctggtacc attttgacgg cgttgattgg 600
gacgaaagcc gaaaattgag ccgcatttac aaattccgcg gcatcggcaa agcgtgggat 660
tgggaagtag acacggaaaa cggaaactat gactacttaa tgtatgccga cttggacatg 720
gaccatcctg aagtggttac ggaactgaaa aactggggca aatggtatgt caacacaacg 780
aacattgatg ggttccggct tgatgccgtc aagcatatta agttcagttt ttttcctgat 840
tggttgtcgt atgtgcgttc tcagactggc aagccgctat ttaccgtcgg ggaatattgg 900
agctatgaca tcaacaagtt gcacaattac attacgaaaa caaacggaac gatgtctttg 960
tttgatgccc cgttacacaa caaattttat accgcttcca aatcaggggg cgcatttgat 1020
atgcgcacgt taatgaccaa tactctcatg aaagatcaac cgacattggc cgtcaccttc 1080
gttgataatc atgacaccga acccggccaa gcgctgcagt catgggtcga cccatggttc 1140
aaaccgttgg cttacgcctt tattctaact cggcaggaag gatacccgtg cgtcttttat 1200
ggtgactatt atggcatccc acaatataac attccttcgc tgaaaagcaa aatcgatccg 1260
ctcctcatcg cgcgcaggga ttatgcttac ggaacgcaac atgattatct tgatcactcc 1320
gacatcatcg ggtggacaag ggaaggcgtc actgaaaaac caggatccgg actggccgca 1380
ctgatcaccg atgggccggg aggaagcaaa tggatgtacg ttggcaaaca acacgccgga 1440
aaagtgttct atgaccttac cggcaaccgg agtgacaccg tcaccatcaa cagtgatgga 1500
tggggagaat tcaaagtcaa tggcggttcg gtttcggttt gggttcctag aaaaacgacc 1560
gtctctacca tcgcttggcc gatcacaacc cgaccgtgga ctggtgaatt cgtccgttgg 1620
accgaaccac ggttggtggc atggccttga 1650

<210> 92
<211> 549
<212> PRT
<213> Bacterial

<400> 92
Val Leu Thr Phe His Arg Ile Ile Arg Lys Gly Trp Met Phe Leu Leu
1 5 10 15
Ala Phe Leu Leu Thr Ala Ser Leu Phe Cys Pro Thr Gly Gln Pro Ala
20 25 30
Lys Ala Ala Ala Pro Phe Asn Gly Thr Met Met Gln Tyr Phe Glu Trp
35 40 45
Tyr Leu Pro Asp Asp Gly Thr Leu Trp Thr Lys Val Ala Asn Glu Ala
50 55 60
Asn Asn Leu Ser Ser Leu Gly Ile Thr Ala Leu Trp Leu Pro Pro Ala
65 70 75 80
Tyr Lys Gly Thr Ser Arg Ser Asp Val Gly Tyr Gly Val Tyr Asp Leu
85 90 95
Tyr Asp Leu Gly Glu Phe Asn Gln Lys Gly Thr Val Arg Thr Lys Tyr
100 105 110
Gly Thr Lys Ala Gln Tyr Leu Gln Ala Ile Gln Ala Ala His Ala Ala
115 120 125
Gly Met Gln Val Tyr Ala Asp Val Val Phe Asp His Lys Gly Gly Ala
130 135 140
Asp Gly Thr Glu Trp Val Asp Ala Val Glu Val Asn Pro Ser Asp Arg
145 150 155 160
Asn Gln Glu Ile Ser Gly Thr Tyr Gln Ile Gln Ala Trp Thr Lys Phe

```
                165                 170                 175
Asp Phe Pro Gly Arg Gly Asn Thr Tyr Ser Ser Phe Lys Trp Arg Trp
            180                 185                 190
Tyr His Phe Asp Gly Val Asp Trp Asp Glu Ser Arg Lys Leu Ser Arg
        195                 200                 205
Ile Tyr Lys Phe Arg Gly Ile Gly Lys Ala Trp Asp Trp Glu Val Asp
    210                 215                 220
Thr Glu Asn Gly Asn Tyr Asp Tyr Leu Met Tyr Ala Asp Leu Asp Met
225                 230                 235                 240
Asp His Pro Glu Val Val Thr Glu Leu Lys Asn Trp Gly Lys Trp Tyr
                245                 250                 255
Val Asn Thr Thr Asn Ile Asp Gly Phe Arg Leu Asp Ala Val Lys His
            260                 265                 270
Ile Lys Phe Ser Phe Phe Pro Asp Trp Leu Ser Tyr Val Arg Ser Gln
        275                 280                 285
Thr Gly Lys Pro Leu Phe Thr Val Gly Glu Tyr Trp Ser Tyr Asp Ile
    290                 295                 300
Asn Lys Leu His Asn Tyr Ile Thr Lys Thr Asn Gly Thr Met Ser Leu
305                 310                 315                 320
Phe Asp Ala Pro Leu His Asn Lys Phe Tyr Thr Ala Ser Lys Ser Gly
                325                 330                 335
Gly Ala Phe Asp Met Arg Thr Leu Met Thr Asn Thr Leu Met Lys Asp
            340                 345                 350
Gln Pro Thr Leu Ala Val Thr Phe Val Asp Asn His Asp Thr Glu Pro
        355                 360                 365
Gly Gln Ala Leu Gln Ser Trp Val Asp Pro Trp Phe Lys Pro Leu Ala
    370                 375                 380
Tyr Ala Phe Ile Leu Thr Arg Gln Glu Gly Tyr Pro Cys Val Phe Tyr
385                 390                 395                 400
Gly Asp Tyr Tyr Gly Ile Pro Gln Tyr Asn Ile Pro Ser Leu Lys Ser
                405                 410                 415
Lys Ile Asp Pro Leu Leu Ile Ala Arg Arg Asp Tyr Ala Tyr Gly Thr
            420                 425                 430
Gln His Asp Tyr Leu Asp His Ser Asp Ile Ile Gly Trp Thr Arg Glu
        435                 440                 445
Gly Val Thr Glu Lys Pro Gly Ser Gly Leu Ala Ala Leu Ile Thr Asp
    450                 455                 460
Gly Pro Gly Gly Ser Lys Trp Met Tyr Val Gly Lys Gln His Ala Gly
465                 470                 475                 480
Lys Val Phe Tyr Asp Leu Thr Gly Asn Arg Ser Asp Thr Val Thr Ile
                485                 490                 495
Asn Ser Asp Gly Trp Gly Glu Phe Lys Val Asn Gly Gly Ser Val Ser
            500                 505                 510
Val Trp Val Pro Arg Lys Thr Thr Val Ser Thr Ile Ala Trp Pro Ile
        515                 520                 525
Thr Thr Arg Pro Trp Thr Gly Glu Phe Val Arg Trp Thr Glu Pro Arg
    530                 535                 540
Leu Val Ala Trp Pro
545
```

<210> 93
<211> 1545
<212> DNA
<213> Bacterial

<400> 93

```
atgaaatcgt ttgcattcat gcctatcctt ttttatgcaa acgatttcat cagtgaaagg      60
gaaggaggag gaaaaatggg gaagaatatg agaagaagat tcacgtattt ttcaatcttc     120
ttattgttcg ttcagctgtt ttcatttagt gcaaccgcta gcgccaatgg aacggtgaac     180
agtagtcctg tggttaatgg aaacgaagtc acgtttctat atggaggaac aggaaacgag     240
cagtctgtgt tactggcagg ctcctttaat gattggcaga aagatggtga caagaagatt     300
gcactaacaa aaggcgacaa taacgtctgg tctgtcacgc aaacacttca agatgggaca     360
tatacgtata agtttgttgt agatggtcaa tgggtggcgg atccgcttaa cccgaatcaa     420
gtagacgacg gttacggcgg ccgtaatagt gtcgttgttg tcgggacacc ggtgcaacaa     480
gaacggacag tgacgcttgt tggtaactta caagacgaat taggtcatac gagcgaatgg     540
gatccgaaag cgacagctac agtgatgaaa aaggaaggga acgggttata tacgtttaca     600
ggtacacttc cagccggaac gtacgagtat aaaattgcga ttaatggcag ctgggacgaa     660
aactatggtg tcggcggccg cgatggcggg aatattaagc tgctattaaa tgaacaaaca     720
```

acggttacat tttattacaa cgacagaacg catgcgattg cggattcgac ttggtatgca 780
ccaattctaa aagaaaagca gccgcggctc gttggaacga ttttaccagc tattggttat 840
gaaacagacg tgaacggttg gacgccgcaa acatcaacgg cgttgttgtc agatgatgat 900
tttgattcca tttatacgtt taaggcgcgt gtgccaaaag ggacatatga atataaagta 960
gttcttggga atgattggac atatgaaaat tatccacaag ataatgccaa attaaatgtg 1020
cttgaagaaa cgacaattac ctttttcttt aacgcgaaaa cgaaagtagt gtataccgat 1080
tacaatccaa gcggttcgga tggtatcgtc caaaaagacc gtttgaagca taatacgtgg 1140
gattcgttgt atcgccaacc gtttggtgcg gtgaaagctg ggacagaagt gacccttcgt 1200
ttatcagcga aaaaaggtga tttgacaaaa gcggatgtat atgtaaaaaa tacgacaacc 1260
ggcacagcga aactatattc gatgaaaaaa gccggtgttc ttggcgaaga agaatattgg 1320
gaagcgacat tcacaccgga tgtgaaagga gtatacggtt ataaatttat tgcggtagat 1380
gctggaacga aagcagaata cggggaagat acacaagaag ggcagtgggg aaaagcagta 1440
gataaaaatg cagagctgtt ccaattaacg gtgtacgacc catcctacca aacaccggat 1500
tggatgaaag aagcagttgt atatcaaatt ttccctgatc caaag 1545

<210> 94
<211> 515
<212> PRT
<213> Bacterial

<400> 94
Met Lys Ser Phe Ala Phe Met Pro Ile Leu Phe Tyr Ala Asn Asp Phe
 1   5   10   15
Ile Ser Glu Arg Glu Gly Gly Gly Lys Met Gly Lys Asn Met Arg Arg
            20   25   30
Arg Phe Thr Tyr Phe Ser Ile Phe Leu Leu Phe Val Gln Leu Phe Ser
        35   40   45
Phe Ser Ala Thr Ala Ser Ala Asn Gly Thr Val Asn Ser Ser Pro Val
    50   55   60
Val Asn Gly Asn Glu Val Thr Phe Leu Tyr Gly Gly Thr Gly Asn Glu
65   70   75   80
Gln Ser Val Leu Leu Ala Gly Ser Phe Asn Asp Trp Gln Lys Asp Gly
                85   90   95
Asp Lys Lys Ile Ala Leu Thr Lys Gly Asp Asn Asn Val Trp Ser Val
            100   105   110
Thr Gln Thr Leu Gln Asp Gly Thr Tyr Thr Tyr Lys Phe Val Val Asp
        115   120   125
Gly Gln Trp Val Ala Asp Pro Leu Asn Pro Asn Gln Val Asp Asp Gly
    130   135   140
Tyr Gly Gly Arg Asn Ser Val Val Val Val Gly Thr Pro Val Gln Gln
145   150   155   160
Glu Arg Thr Val Thr Leu Val Gly Asn Leu Gln Asp Glu Leu Gly His
                165   170   175
Thr Ser Glu Trp Asp Pro Lys Ala Thr Ala Thr Val Met Lys Lys Glu
            180   185   190
Gly Asn Gly Leu Tyr Thr Phe Thr Gly Thr Leu Pro Ala Gly Thr Tyr
        195   200   205
Glu Tyr Lys Ile Ala Ile Asn Gly Ser Trp Asp Glu Asn Tyr Gly Val
    210   215   220
Gly Gly Arg Asp Gly Gly Asn Ile Lys Leu Leu Leu Asn Glu Gln Thr
225   230   235   240
Thr Val Thr Phe Tyr Tyr Asn Asp Arg Thr His Ala Ile Ala Asp Ser
                245   250   255
Thr Trp Tyr Ala Pro Ile Leu Lys Glu Lys Gln Pro Arg Leu Val Gly
            260   265   270
Thr Ile Leu Pro Ala Ile Gly Tyr Glu Thr Asp Val Asn Gly Trp Thr
        275   280   285
Pro Gln Thr Ser Thr Ala Leu Leu Ser Asp Asp Asp Phe Asp Ser Ile
    290   295   300
Tyr Thr Phe Lys Ala Arg Val Pro Lys Gly Thr Tyr Glu Tyr Lys Val
305   310   315   320
Val Leu Gly Asn Asp Trp Thr Tyr Glu Asn Tyr Pro Gln Asp Asn Ala
                325   330   335
Lys Leu Asn Val Leu Glu Glu Thr Thr Ile Thr Phe Phe Phe Asn Ala
            340   345   350
Lys Thr Lys Val Val Tyr Thr Asp Tyr Asn Pro Ser Gly Ser Asp Gly
        355   360   365
Ile Val Gln Lys Asp Arg Leu Lys His Asn Thr Trp Asp Ser Leu Tyr

```
    370                 375                 380
Arg Gln Pro Phe Gly Ala Val Lys Ala Gly Thr Glu Val Thr Leu Arg
385                 390                 395                 400
Leu Ser Ala Lys Lys Gly Asp Leu Thr Lys Ala Asp Val Tyr Val Lys
                405                 410                 415
Asn Thr Thr Thr Gly Thr Ala Lys Leu Tyr Ser Met Lys Lys Ala Gly
            420                 425                 430
Val Leu Gly Glu Glu Glu Tyr Trp Glu Ala Thr Phe Thr Pro Asp Val
        435                 440                 445
Lys Gly Val Tyr Gly Tyr Lys Phe Ile Ala Val Asp Ala Gly Thr Lys
    450                 455                 460
Ala Glu Tyr Gly Glu Asp Thr Gln Glu Gly Gln Trp Gly Lys Ala Val
465                 470                 475                 480
Asp Lys Asn Ala Glu Leu Phe Gln Leu Thr Val Tyr Asp Pro Ser Tyr
                485                 490                 495
Gln Thr Pro Asp Trp Met Lys Glu Ala Val Val Tyr Gln Ile Phe Pro
            500                 505                 510
Asp Pro Lys
        515
```

<210> 95
<211> 1524
<212> DNA
<213> Bacterial

<400> 95

```
atgtatacac tattcatccg ttcatatttt gatactgatg gtgatggtgt aggagacttt      60
agtggagttg ctgaaaaggt agattatcta aaatctcttg gagtagatac agtctggttt     120
ttaccattta ataaaagtaa atcttatcat ggatatgatg ttgaagatta ctatgatgta     180
gaaccagatt atggaacact acaagatctt gataatatga taaaagttct aaatgaaaat     240
ggaataaagg tagtaatgga tcttgttgtt aatcatacgt cggatacaca tccatggttt     300
cttgatgcag ttgaaaatac tactaattct ccatattgga actattacat tatgagcttg     360
gatgagcctc aaaataagaa tcattggcat tataaggtta attcaaaagg acaaactgtg     420
tggtattttg gattgtttga ttcatcaatg ccggacctta attacgacaa ccctaaagta     480
atggatgaag tgaaaaaaat aatagatttt tgggcagata tgggagtaga tggatttaga     540
ttagatgcag caaaacatta ttatggattt gactggagcg atggaattga acagtcagca     600
agcgttgcaa aagagataga agactatata aaagataaac taggggaaaa tgcaatagtt     660
gtgagtgagg tttacgatgg agattcaaat gttcttttaa aatttgctcc aatgcctgtg     720
tttaatttta gttttatgta caatttgaga ggaaattttg aagggagaga taacttaatt     780
tcagactcta ttagttgggt tgattcctcg ttgtataatt taaatgtttt tcattttcca     840
tttattgata gtcatgatct tgacagattt atttctgagc ttgtagatag taaatatcag     900
ggagatgtaa tatctgccac aaaacaatat ttgctagtta atgctttact actctcatta     960
acaggcatgc caactattta ctatggtgat gaaataggac ttaggggatg gaagtggcat    1020
tcagaaccat gggatatacc tgtgcgtgag ccaatgcaat ggtataagga tcaaaaaggg    1080
aacggtcaaa cttattggac aaaagagttt tacgaaggta ttactgaagg aagtgctaat    1140
gaagatggag caatatacga tgatccagat gatggagtat ctgtagaaga acaagaaaat    1200
ggatattcta ttttaaactt ttttaaagaa tttatcaact tacgaaaaga ttatccggca    1260
cttgcttttg gaagtactac gattgagaga gattggaaaa acttgtatgt tttgaaaaag    1320
tcgtataact tccaggatgt tcttgtatta attaaccttg atccaacgta ttcaaataca    1380
tacgaagttc cagaagggta taaatgggtg tggtatgcat tttttgatgg tgacaactat    1440
gaatttggag caaaagatga aatgatttta cagaatacaa gttggacgat aaatccaagg    1500
caaatttata tatttgtaaa gtaa                                           1524
```

<210> 96
<211> 507
<212> PRT
<213> Bacterial

<400> 96

```
Met Tyr Thr Leu Phe Ile Arg Ser Tyr Phe Asp Thr Asp Gly Asp Gly
 1               5                  10                  15
Val Gly Asp Phe Ser Gly Val Ala Glu Lys Val Asp Tyr Leu Lys Ser
            20                  25                  30
Leu Gly Val Asp Thr Val Trp Phe Leu Pro Phe Asn Lys Ser Lys Ser
        35                  40                  45
Tyr His Gly Tyr Asp Val Glu Asp Tyr Tyr Asp Val Glu Pro Asp Tyr
    50                  55                  60
Gly Thr Leu Gln Asp Leu Asp Asn Met Ile Lys Val Leu Asn Glu Asn
```

65 70 75 80
Gly Ile Lys Val Val Met Asp Leu Val Val Asn His Thr Ser Asp Thr
85 90 95
His Pro Trp Phe Leu Asp Ala Val Glu Asn Thr Thr Asn Ser Pro Tyr
100 105 110
Trp Asn Tyr Tyr Ile Met Ser Leu Asp Glu Pro Gln Asn Lys Asn His
115 120 125
Trp His Tyr Lys Val Asn Ser Lys Gly Gln Thr Val Trp Tyr Phe Gly
130 135 140
Leu Phe Asp Ser Ser Met Pro Asp Leu Asn Tyr Asp Asn Pro Lys Val
145 150 155 160
Met Asp Glu Val Lys Lys Ile Ile Asp Phe Trp Ala Asp Met Gly Val
165 170 175
Asp Gly Phe Arg Leu Asp Ala Ala Lys His Tyr Tyr Gly Phe Asp Trp
180 185 190
Ser Asp Gly Ile Glu Gln Ser Ala Ser Val Ala Lys Glu Ile Glu Asp
195 200 205
Tyr Ile Lys Asp Lys Leu Gly Glu Asn Ala Ile Val Val Ser Glu Val
210 215 220
Tyr Asp Gly Asp Ser Asn Val Leu Leu Lys Phe Ala Pro Met Pro Val
225 230 235 240
Phe Asn Phe Ser Phe Met Tyr Asn Leu Arg Gly Asn Phe Glu Gly Arg
245 250 255
Asp Asn Leu Ile Ser Asp Ser Ile Ser Trp Val Asp Ser Ser Leu Tyr
260 265 270
Asn Leu Asn Val Phe His Phe Pro Phe Ile Asp Ser His Asp Leu Asp
275 280 285
Arg Phe Ile Ser Glu Leu Val Asp Ser Lys Tyr Gln Gly Asp Val Ile
290 295 300
Ser Ala Thr Lys Gln Tyr Leu Leu Val Asn Ala Leu Leu Leu Ser Leu
305 310 315 320
Thr Gly Met Pro Thr Ile Tyr Tyr Gly Asp Glu Ile Gly Leu Arg Gly
325 330 335
Trp Lys Trp His Ser Glu Pro Trp Asp Ile Pro Val Arg Glu Pro Met
340 345 350
Gln Trp Tyr Lys Asp Gln Lys Gly Asn Gly Gln Thr Tyr Trp Thr Lys
355 360 365
Glu Phe Tyr Glu Gly Ile Thr Glu Gly Ser Ala Asn Glu Asp Gly Ala
370 375 380
Ile Tyr Asp Asp Pro Asp Asp Gly Val Ser Val Glu Glu Gln Glu Asn
385 390 395 400
Gly Tyr Ser Ile Leu Asn Phe Phe Lys Glu Phe Ile Asn Leu Arg Lys
405 410 415
Asp Tyr Pro Ala Leu Ala Phe Gly Ser Thr Thr Ile Glu Arg Asp Trp
420 425 430
Lys Asn Leu Tyr Val Leu Lys Lys Ser Tyr Asn Phe Gln Asp Val Leu
435 440 445
Val Leu Ile Asn Leu Asp Pro Thr Tyr Ser Asn Thr Tyr Glu Val Pro
450 455 460
Glu Gly Tyr Lys Trp Val Trp Tyr Ala Phe Phe Asp Gly Asp Asn Tyr
465 470 475 480
Glu Phe Gly Ala Lys Asp Glu Met Ile Leu Gln Asn Thr Ser Trp Thr
485 490 495
Ile Asn Pro Arg Gln Ile Tyr Ile Phe Val Lys
500 505

<210> 97
<211> 2931
<212> DNA
<213> Bacterial

<400> 97
atgaggaaga agatgtcgca ttcaagattt acttttcttt tgatcttagc acttttttatt 60
ttcttctccg gttgtatttc agaagttaaa agcgaaagcc agctactaaa ttcaaagcaa 120
aaggtccttg taaaagtaaa tgttaatacg ccatttattg agaatgctac tactaatacg 180
tggagtgttt caaaagaatc ttttattgat tatcttagta aagtgattat tactgttaag 240
gatgtaaatg atcagattgt atttactaag gaaacaacga acaaaacaaa tatttatttt 300
gaaattgaac ttcttcctgg aacttataca tttgaggtaa aaggatatga ggaagattta 360

```
gttatatttt caggggaaaa agttaatcag atcatagatg agaaaaataa tattgttaat     420
gtcgaaactt tttttgttaa tggaatagtt aggacaataa ttgaagttga cgatattatt     480
tataaaaatt atgatattac atcggcaacg ttgatcttca aaaaagatac agcacaagaa     540
gattatgaag aggtacctgt aacacttaca ggtacttcca ctttaattaa taaagaatta     600
tatcctggta tgtggactgt aaaatttgaa gttgatctta aatcaaagga tgcaagtatg     660
ttaccagaaa aagttcatct tgaaaatgaa tttagcatag aagtgcttcc agcaaagaca     720
aaaagtttaa catttaatgt agtctttgat acagaggtta atgaaccgaa attagtagtt     780
gtatttccgc aaattgagtt gccttttgtg gatcctgtaa caaatttaag tggagagata     840
aatgaattag aagggaatct ttcaatgaat tgggactatt cagatccaaa tgcagaattt     900
tatgtgtata aagaattaga ggaacaagga gaatatttgt atgaatttgt tggaaaaaca     960
cgcgagaaaa gttatacaat agaaaatttt accaagcaag aattcgataa atttagtgga    1020
atcgctatta atgtttatgc caacggtaaa gagagtggat tagttgttct aaaaaaagaa    1080
aatattaaac ttatagattt agaaagtgtt gacagtataa gtgctactta taacgttgat    1140
acgaatgagc ttaagttgga ttggaattat accaattcaa gtgttacttt tgaagttttg    1200
aaaaaaggta taaatagcaa tgaatacgaa ataatttctc aactaacaca aaattctttt    1260
tcaacagaat tcacaggcag gcaattttgg gatcttgaga aaattgcgat tagagtagtt    1320
gctaatggat ttgaaagtaa gattaatgag atttcaagag atgatataac tataacatca    1380
ttgaatcttc ctcttacatc gtctactatg tatacactat tcatccgttc atattttgat    1440
actgatggtg atggtgtagg agactttagt ggagttgctg aaaaggtaga ttatctaaaa    1500
tctcttggag tagatacagt ctggttttta ccatttaata aaagtaaatc ttatcatgga    1560
tatgatgttg aagattacta tgatgtagaa ccagattatg gaacactaca agatcttgat    1620
aatatgataa aagttctaaa tgaaaatgga ataaaggtag taatggatct tgttgttaat    1680
catacgtcgg atacacatcc atggtttctt gatgcagttg aaaatactac taattctcca    1740
tattggaact attacattat gagcttggat gagcctcaaa ataagaatca ttggcattat    1800
aaggttaatt caaaaggaca aactgtgtgg tattttggat tgtttgattc atcaatgccg    1860
gaccttaatt acgacaaccc taaagtaatg gatgaagtga aaaaaataat agatttttgg    1920
gcagatatgg gagtagatgg atttagatta gatgcagcaa aacattatta tggatttgac    1980
tggagcgatg gaattgaaca gtcagcaagc gttgcaaaag agatagaaga ctatataaaa    2040
gataaactag gggaaaatgc aatagttgtg agtgaggttt acgatggaga ttcaaatgtt    2100
cttttaaaat ttgctccaat gcctgtgttt aattttagtt ttatgtacaa tttgagagga    2160
aattttgaag ggagagataa cttaatttca gactctatta gttgggttga ttcctcgttg    2220
tataatttaa atgtttttca ttttccattt attgatagtc atgatcttga cagatttatt    2280
tctgagcttg tagatagtaa atatcaggga gatgtaatat ctgccacaaa acaatatttg    2340
ctagttaatg ctttactact ctcattaaca ggcatgccaa ctatttacta tggtgatgaa    2400
ataggactta ggggatggaa gtggcattca gaaccatggg atatacctgt gcgtgagcca    2460
atgcaatggt ataaggatca aaaagggaac ggtcaaactt attggacaaa agagttttac    2520
gaaggtatta ctgaaggaag tgctaatgaa gatggagcaa tatacgatga tccagatgat    2580
ggagtatctg tagaagaaca agaaaatgga tattctattt taaacttttt taaagaattt    2640
atcaacttac gaaaagatta tccggcactt gcttttggaa gtactacgat tgagagagat    2700
tggaaaaact tgtatgtttt gaaaaagtcg tataacttcc aggatgttct tgtattaatt    2760
aaccttgatc caacgtattc aaatacatac gaagttccag aagggtataa atgggtgtgg    2820
tatgcatttt ttgatggtga caactatgaa tttggagcaa aagatgaaat gattttacag    2880
aatacaagtt ggacgataaa tccaaggcaa atttatatat ttgtaaagta a              2931

<210> 98
<211> 976
<212> PRT
<213> Bacterial

<400> 98
Met Arg Lys Lys Met Ser His Ser Arg Phe Thr Phe Leu Leu Ile Leu
 1               5                  10                  15
Ala Leu Phe Ile Phe Phe Ser Gly Cys Ile Ser Glu Val Lys Ser Glu
            20                  25                  30
Ser Gln Leu Leu Asn Ser Lys Gln Lys Val Leu Val Lys Val Asn Val
        35                  40                  45
Asn Thr Pro Phe Ile Glu Asn Ala Thr Thr Asn Thr Trp Ser Val Ser
    50                  55                  60
Lys Glu Ser Phe Ile Asp Tyr Leu Ser Lys Val Ile Ile Thr Val Lys
65                  70                  75                  80
Asp Val Asn Asp Gln Ile Val Phe Thr Lys Glu Thr Thr Asn Lys Thr
                85                  90                  95
Asn Ile Tyr Phe Glu Ile Glu Leu Leu Pro Gly Thr Tyr Thr Phe Glu
            100                 105                 110
Val Lys Gly Tyr Glu Glu Asp Leu Val Ile Phe Ser Gly Glu Lys Val
        115                 120                 125
Asn Gln Ile Ile Asp Glu Lys Asn Asn Ile Val Asn Val Glu Thr Phe
    130                 135                 140
```

```
Phe Val Asn Gly Ile Val Arg Thr Ile Ile Glu Val Asp Asp Ile Ile
145                 150                 155                 160
Tyr Lys Asn Tyr Asp Ile Thr Ser Ala Thr Leu Ile Phe Lys Lys Asp
                165                 170                 175
Thr Ala Gln Glu Asp Tyr Glu Glu Val Pro Val Thr Leu Thr Gly Thr
            180                 185                 190
Ser Thr Leu Ile Asn Lys Glu Leu Tyr Pro Gly Met Trp Thr Val Lys
        195                 200                 205
Phe Glu Val Asp Leu Lys Ser Lys Asp Ala Ser Met Leu Pro Glu Lys
    210                 215                 220
Val His Leu Glu Asn Glu Phe Ser Ile Glu Val Leu Pro Ala Lys Thr
225                 230                 235                 240
Lys Ser Leu Thr Phe Asn Val Val Phe Asp Thr Glu Val Asn Glu Pro
                245                 250                 255
Lys Leu Val Val Val Phe Pro Gln Ile Glu Leu Pro Phe Val Asp Pro
            260                 265                 270
Val Thr Asn Leu Ser Gly Glu Ile Asn Glu Leu Glu Gly Asn Leu Ser
        275                 280                 285
Met Asn Trp Asp Tyr Ser Asp Pro Asn Ala Glu Phe Tyr Val Tyr Lys
    290                 295                 300
Glu Leu Glu Glu Gln Gly Glu Tyr Leu Tyr Glu Phe Val Gly Lys Thr
305                 310                 315                 320
Arg Glu Lys Ser Tyr Thr Ile Glu Asn Phe Thr Lys Gln Glu Phe Asp
                325                 330                 335
Lys Phe Ser Gly Ile Ala Ile Asn Val Tyr Ala Asn Gly Lys Glu Ser
            340                 345                 350
Gly Leu Val Val Leu Lys Lys Glu Asn Ile Lys Leu Ile Asp Leu Glu
        355                 360                 365
Ser Val Asp Ser Ile Ser Ala Thr Tyr Asn Val Asp Thr Asn Glu Leu
    370                 375                 380
Lys Leu Asp Trp Asn Tyr Thr Asn Ser Ser Val Thr Phe Glu Val Leu
385                 390                 395                 400
Lys Lys Gly Ile Asn Ser Asn Glu Tyr Glu Ile Ile Ser Gln Leu Thr
                405                 410                 415
Gln Asn Ser Phe Ser Thr Glu Phe Thr Gly Arg Gln Phe Trp Asp Leu
            420                 425                 430
Glu Lys Ile Ala Ile Arg Val Val Ala Asn Gly Phe Glu Ser Lys Ile
        435                 440                 445
Asn Glu Ile Ser Arg Asp Asp Ile Thr Ile Thr Ser Leu Asn Leu Pro
    450                 455                 460
Leu Thr Ser Ser Thr Met Tyr Thr Leu Phe Ile Arg Ser Tyr Phe Asp
465                 470                 475                 480
Thr Asp Gly Asp Gly Val Gly Asp Phe Ser Gly Val Ala Glu Lys Val
                485                 490                 495
Asp Tyr Leu Lys Ser Leu Gly Val Asp Thr Val Trp Phe Leu Pro Phe
            500                 505                 510
Asn Lys Ser Lys Ser Tyr His Gly Tyr Asp Val Glu Asp Tyr Tyr Asp
        515                 520                 525
Val Glu Pro Asp Tyr Gly Thr Leu Gln Asp Leu Asp Asn Met Ile Lys
    530                 535                 540
Val Leu Asn Glu Asn Gly Ile Lys Val Val Met Asp Leu Val Val Asn
545                 550                 555                 560
His Thr Ser Asp Thr His Pro Trp Phe Leu Asp Ala Val Glu Asn Thr
                565                 570                 575
Thr Asn Ser Pro Tyr Trp Asn Tyr Tyr Ile Met Ser Leu Asp Glu Pro
            580                 585                 590
Gln Asn Lys Asn His Trp His Tyr Lys Val Asn Ser Lys Gly Gln Thr
        595                 600                 605
Val Trp Tyr Phe Gly Leu Phe Asp Ser Ser Met Pro Asp Leu Asn Tyr
    610                 615                 620
Asp Asn Pro Lys Val Met Asp Glu Val Lys Lys Ile Ile Asp Phe Trp
625                 630                 635                 640
Ala Asp Met Gly Val Asp Gly Phe Arg Leu Asp Ala Ala Lys His Tyr
                645                 650                 655
Tyr Gly Phe Asp Trp Ser Asp Gly Ile Glu Gln Ser Ala Ser Val Ala
            660                 665                 670
Lys Glu Ile Glu Asp Tyr Ile Lys Asp Lys Leu Gly Glu Asn Ala Ile
        675                 680                 685
```

```
Val Val Ser Glu Val Tyr Asp Gly Asp Ser Asn Val Leu Leu Lys Phe
    690                 695                 700
Ala Pro Met Pro Val Phe Asn Phe Ser Phe Met Tyr Asn Leu Arg Gly
705                 710                 715                 720
Asn Phe Glu Gly Arg Asp Asn Leu Ile Ser Asp Ser Ile Ser Trp Val
                725                 730                 735
Asp Ser Ser Leu Tyr Asn Leu Asn Val Phe His Phe Pro Phe Ile Asp
            740                 745                 750
Ser His Asp Leu Asp Arg Phe Ile Ser Glu Leu Val Asp Ser Lys Tyr
        755                 760                 765
Gln Gly Asp Val Ile Ser Ala Thr Lys Gln Tyr Leu Leu Val Asn Ala
    770                 775                 780
Leu Leu Leu Ser Leu Thr Gly Met Pro Thr Ile Tyr Tyr Gly Asp Glu
785                 790                 795                 800
Ile Gly Leu Arg Gly Trp Lys Trp His Ser Glu Pro Trp Asp Ile Pro
                805                 810                 815
Val Arg Glu Pro Met Gln Trp Tyr Lys Asp Gln Lys Gly Asn Gly Gln
            820                 825                 830
Thr Tyr Trp Thr Lys Glu Phe Tyr Glu Gly Ile Thr Glu Gly Ser Ala
        835                 840                 845
Asn Glu Asp Gly Ala Ile Tyr Asp Asp Pro Asp Asp Gly Val Ser Val
    850                 855                 860
Glu Glu Gln Glu Asn Gly Tyr Ser Ile Leu Asn Phe Phe Lys Glu Phe
865                 870                 875                 880
Ile Asn Leu Arg Lys Asp Tyr Pro Ala Leu Ala Phe Gly Ser Thr Thr
                885                 890                 895
Ile Glu Arg Asp Trp Lys Asn Leu Tyr Val Leu Lys Lys Ser Tyr Asn
            900                 905                 910
Phe Gln Asp Val Leu Val Leu Ile Asn Leu Asp Pro Thr Tyr Ser Asn
        915                 920                 925
Thr Tyr Glu Val Pro Glu Gly Tyr Lys Trp Val Trp Tyr Ala Phe Phe
    930                 935                 940
Asp Gly Asp Asn Tyr Glu Phe Gly Ala Lys Asp Glu Met Ile Leu Gln
945                 950                 955                 960
Asn Thr Ser Trp Thr Ile Asn Pro Arg Gln Ile Tyr Ile Phe Val Lys
                965                 970                 975
```

<210> 99
<211> 1536
<212> DNA
<213> Bacterial

<400> 99

```
atgtacacac tcttcatccg ctctttttac gatacaaaca acgacggtgt aggtgactac      60
aacggtgttg cccaaaaagt agactatctc aaaacgcttg gagtggatac agtttggttc     120
ttgccgttca acaaagcaaa atcgtaccac ggttacgatg ttgaagacta ctacgatgta     180
gaacctgact atggaacata cgcacaactt gaaaatatga taaagacact caatcagaac     240
ggaattcgtg ttgttatgga cttggttgtg aaccacactt ccgatacaca ctcgtggttt     300
ctggatgccg ttgagaacac aacgaattcg aaatattgga gctactacat aatgacactt     360
gaaaatagag acggttggaa tcactggcat tggaagataa actcaaaagg gcaaaaagtt     420
tactacttcg gactgtttga ctcatcaatg cccgatttga atttcgacaa tccacaagtg     480
atgaacgaaa tcaagagaat aatcgatttc tggataacag ttggtgtgga tggtttcaga     540
cttgatgcac caaagcacta caaaggctgg gattgggacg acggcatttc aggttcagca     600
gcaatcgcga gggaaataga aagttacatc aggagcaagt taggaaacga tgcgatagtt     660
gtcggggaag tgtacgatgg aaatccatcg gttctttcac aatttgcacc gatgccggcg     720
ttcaacttca cattcatgta tggaataaca ggcaaccatg aggggaaaga taacctgctg     780
ggagaaacaa tttcatgggt taatggagcg agttattatc tcaacgtaaa acatttcccg     840
ttcatagaca atcacgattt gaacagatgg atatcgatac ttatcgacca aaagtatagt     900
ggaaacacac aagttggtac gaagcagtat attttaacaa atgcgctctt gctttcctta     960
aacggtatgc ctgttattta ttatgggaat gaaataggct tgagaggatg gaaatgggga    1020
caagacccgt gggatttgcc ggtgagagag ccgatgcagt ggtacgcaag tcaaagtgga    1080
gctgggcaga catggtggac aaagcctgtc taccagcaaa aaggaatcac atttggaaat    1140
gcaaacgtcg atggtgcgat gtacgatgat ccaaatgatg gggtttcagt agaagagcag    1200
atgaatggtt acacgataaa taacttcttt aaacaattca taaccctgag gaagacatat    1260
ccggctctat cgaaaggttc gataacgata gaacgcgact ggaagaacct gtacgttatc    1320
aaacgagtct acggaaatca ggaagtgctt gtattgataa acttagaccc aacttggccg    1380
aacaattaca cgttaccagg tggatacagg tgggtctggt atgcgttctt taatgggagt    1440
ttgtttgaat ttggcaataa aaacgaatca ccactgagcc aagataccaa ctggacagtc    1500
```

aatccaaggc aagtgtatgt gtttgtgaag gactaa 1536

<210> 100
<211> 511
<212> PRT
<213> Bacterial

<400> 100

```
Met Tyr Thr Leu Phe Ile Arg Ser Phe Tyr Asp Thr Asn Asn Asp Gly
 1               5                  10                  15
Val Gly Asp Tyr Asn Gly Val Ala Gln Lys Val Asp Tyr Leu Lys Thr
            20                  25                  30
Leu Gly Val Asp Thr Val Trp Phe Leu Pro Phe Asn Lys Ala Lys Ser
        35                  40                  45
Tyr His Gly Tyr Asp Val Glu Asp Tyr Tyr Asp Val Glu Pro Asp Tyr
    50                  55                  60
Gly Thr Tyr Ala Gln Leu Glu Asn Met Ile Lys Thr Leu Asn Gln Asn
65                  70                  75                  80
Gly Ile Arg Val Val Met Asp Leu Val Val Asn His Thr Ser Asp Thr
                85                  90                  95
His Ser Trp Phe Leu Asp Ala Val Glu Asn Thr Thr Asn Ser Lys Tyr
            100                 105                 110
Trp Ser Tyr Tyr Ile Met Thr Leu Glu Asn Arg Asp Gly Trp Asn His
        115                 120                 125
Trp His Trp Lys Ile Asn Ser Lys Gly Gln Lys Val Tyr Tyr Phe Gly
    130                 135                 140
Leu Phe Asp Ser Ser Met Pro Asp Leu Asn Phe Asp Asn Pro Gln Val
145                 150                 155                 160
Met Asn Glu Ile Lys Arg Ile Ile Asp Phe Trp Ile Thr Val Gly Val
                165                 170                 175
Asp Gly Phe Arg Leu Asp Ala Pro Lys His Tyr Lys Gly Trp Asp Trp
            180                 185                 190
Asp Asp Gly Ile Ser Gly Ser Ala Ala Ile Ala Arg Glu Ile Glu Ser
        195                 200                 205
Tyr Ile Arg Ser Lys Leu Gly Asn Asp Ala Ile Val Val Gly Glu Val
    210                 215                 220
Tyr Asp Gly Asn Pro Ser Val Leu Ser Gln Phe Ala Pro Met Pro Ala
225                 230                 235                 240
Phe Asn Phe Thr Phe Met Tyr Gly Ile Thr Gly Asn His Glu Gly Lys
                245                 250                 255
Asp Asn Leu Leu Gly Glu Thr Ile Ser Trp Val Asn Gly Ala Ser Tyr
            260                 265                 270
Tyr Leu Asn Val Lys His Phe Pro Phe Ile Asp Asn His Asp Leu Asn
        275                 280                 285
Arg Trp Ile Ser Ile Leu Ile Asp Gln Lys Tyr Ser Gly Asn Thr Gln
    290                 295                 300
Val Gly Thr Lys Gln Tyr Ile Leu Thr Asn Ala Leu Leu Leu Ser Leu
305                 310                 315                 320
Asn Gly Met Pro Val Ile Tyr Tyr Gly Asn Glu Ile Gly Leu Arg Gly
                325                 330                 335
Trp Lys Trp Gly Gln Asp Pro Trp Asp Leu Pro Val Arg Glu Pro Met
            340                 345                 350
Gln Trp Tyr Ala Ser Gln Ser Gly Ala Gly Gln Thr Trp Trp Thr Lys
        355                 360                 365
Pro Val Tyr Gln Gln Lys Gly Ile Thr Phe Gly Asn Ala Asn Val Asp
    370                 375                 380
Gly Ala Met Tyr Asp Asp Pro Asn Asp Gly Val Ser Val Glu Glu Gln
385                 390                 395                 400
Met Asn Gly Tyr Thr Ile Asn Asn Phe Phe Lys Gln Phe Ile Thr Leu
                405                 410                 415
Arg Lys Thr Tyr Pro Ala Leu Ser Lys Gly Ser Ile Thr Ile Glu Arg
            420                 425                 430
Asp Trp Lys Asn Leu Tyr Val Ile Lys Arg Val Tyr Gly Asn Gln Glu
        435                 440                 445
Val Leu Val Leu Ile Asn Leu Asp Pro Thr Trp Pro Asn Asn Tyr Thr
    450                 455                 460
Leu Pro Gly Gly Tyr Arg Trp Val Trp Tyr Ala Phe Phe Asn Gly Ser
465                 470                 475                 480
```

Leu Phe Glu Phe Gly Asn Lys Asn Glu Ser Pro Leu Ser Gln Asp Thr
485 490 495
Asn Trp Thr Val Asn Pro Arg Gln Val Tyr Val Phe Val Lys Asp
500 505 510

<210> 101
<211> 2136
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 101
ttgcgattct ttccaaagtt aatatcccct tttccgcaaa acaccagaga gtggcagcga 60
agcgcagtat caagagacac tgaacaatta caaaggaaag taataatgat caatttgaaa 120
aaaaacacca ttagcgccct ggtcgcaggt atggtattag gctttgcatc caacgcaatg 180
gcggttccta gaaccgcttt tgtacacctc tttgaatgga aatgggaaga tgttgcacag 240
gagtgtgaaa catttctcgg acctaaaggc tttgccgcag tgcaagtctc tccgccaact 300
aaatctcaca acacggatgc atggtggggc cgttatcaac ccgttagtta tgcttttgaa 360
ggacgcagcg gtaatcgcag ccaatttaaa aatatggtgc aacgttgtaa agctgtaggc 420
gtcgatatat acgtagatgc agtgattaac cacatggcag cctacgacag aaatttccct 480
gatgtaccct atagcagtaa tgactttaac tcctgtacag gagatattga ctataataac 540
cgttggcaaa cacagcattg tgatttagtc ggtcttaatg atctaaaaac aggatctgac 600
tacgtccgcc aaaaaatagc ggattatatg aacgacgcaa tcagtatggg tgtagctggt 660
ttccgtattg atgcagccaa acatatacca gcaggtgata tagctgccat taaaggtaaa 720
ttaaatggta atccatacat cttccaagag gtaattggtg catccggcga acctgttcga 780
ccgactgaat acacctttat cggtggtgtc acggaatttc aatttgctcg aaaattgggt 840
ccagccttcc gcaatagtaa tattgcttgg ttaaaagaca ttggcagtca aatggaatta 900
tccagtgctg atgccgtaac atttgtaacg aatcatgatg aagagcgtca taacccgaat 960
ggtcctattt ggcacggcgt tcaaggtaat ggttatgcat tagcaaatat tttcacctta 1020
gcttaccctt acggctatcc aaaaatcatg tcaggatact tcttccacgg tgactttaac 1080
gcagctccac caagcagtgg tatacacaca ggaaatgcgt gtggttttga tggcggagac 1140
tgggtatgcg aacacaaatg gcgcggtatt gctaacatgg ttgccttccg caactataca 1200
gcaagcgaat ggcgtatcag taattggtgg caaaacagta acgaccaaat tgcttttggt 1260
cgcggtggtt taggttttgt tgttattaat aaacgtgcta atggtagcat taatcaaagt 1320
tttgatacgg gaatgcctga tggccaatac tgtaacataa tagaagctaa ctttgatgaa 1380
agcaccggcc aatgtagtgc agctacagat tccaacggtc aagccgttat taccgtcagt 1440
ggtgggcaag ctaactttaa tgtagcaggc gatcatgctg ctgcaattca tgttggcgca 1500
aaaattggtg atcaatgtag tggtgatgat tgcccatgta caggatccga ttgtaataat 1560
gatcctaaac ctgattttgc agtaccagca acatcaattt gtacatcaga aaatttacct 1620
acgctatatt actggggagc acagcctaca gatagcttag cgaatgcagc ttggccaggt 1680
gtcgcaatgc aaacaaatgg cgactttaag tgtcatgatt taggtgtcga actaaccaaa 1740
attaacgcca tctttagtga caatggtgca aataaaacag ctgatctaac tgttactggt 1800
gcaggttgtt ataaagacgg gacttggagc accttacaaa attgtggctt tgaaattacc 1860
ggtgcacaaa ccaatccagt cggtggcgac gaagtctggt acttccgagg tactgctaat 1920
gactggggta aagcacaatt agattatgac gcaactagcg gtttgtatta cacaatacaa 1980
agctttaatg gtgaagaagc acctgcgcgt tttaaaattg ataatggtag ttggactgaa 2040
gcttatccaa cagctgatta ccaagttaca gataacaatt cataccgcat taactttaat 2100
agcgatagca aagcgattac agtaaacgca caataa 2136

<210> 102
<211> 711
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 102
Met Arg Phe Phe Pro Lys Leu Ile Ser Pro Phe Pro Gln Asn Thr Arg
1 5 10 15
Glu Trp Gln Arg Ser Ala Val Ser Arg Asp Thr Glu Gln Leu Gln Arg
20 25 30
Lys Val Ile Met Ile Asn Leu Lys Lys Asn Thr Ile Ser Ala Leu Val
35 40 45
Ala Gly Met Val Leu Gly Phe Ala Ser Asn Ala Met Ala Val Pro Arg
50 55 60

```
Thr Ala Phe Val His Leu Phe Glu Trp Lys Trp Glu Asp Val Ala Gln
65                  70                  75                  80
Glu Cys Glu Thr Phe Leu Gly Pro Lys Gly Phe Ala Ala Val Gln Val
                85                  90                  95
Ser Pro Pro Thr Lys Ser His Asn Thr Asp Ala Trp Trp Gly Arg Tyr
            100                 105                 110
Gln Pro Val Ser Tyr Ala Phe Glu Gly Arg Ser Gly Asn Arg Ser Gln
        115                 120                 125
Phe Lys Asn Met Val Gln Arg Cys Lys Ala Val Gly Val Asp Ile Tyr
    130                 135                 140
Val Asp Ala Val Ile Asn His Met Ala Ala Tyr Asp Arg Asn Phe Pro
145                 150                 155                 160
Asp Val Pro Tyr Ser Ser Asn Asp Phe Asn Ser Cys Thr Gly Asp Ile
                165                 170                 175
Asp Tyr Asn Asn Arg Trp Gln Thr Gln His Cys Asp Leu Val Gly Leu
            180                 185                 190
Asn Asp Leu Lys Thr Gly Ser Asp Tyr Val Arg Gln Lys Ile Ala Asp
        195                 200                 205
Tyr Met Asn Asp Ala Ile Ser Met Gly Val Ala Gly Phe Arg Ile Asp
    210                 215                 220
Ala Ala Lys His Ile Pro Ala Gly Asp Ile Ala Ala Ile Lys Gly Lys
225                 230                 235                 240
Leu Asn Gly Asn Pro Tyr Ile Phe Gln Glu Val Ile Gly Ala Ser Gly
                245                 250                 255
Glu Pro Val Arg Pro Thr Glu Tyr Thr Phe Ile Gly Gly Val Thr Glu
            260                 265                 270
Phe Gln Phe Ala Arg Lys Leu Gly Pro Ala Phe Arg Asn Ser Asn Ile
        275                 280                 285
Ala Trp Leu Lys Asp Ile Gly Ser Gln Met Glu Leu Ser Ser Ala Asp
    290                 295                 300
Ala Val Thr Phe Val Thr Asn His Asp Glu Glu Arg His Asn Pro Asn
305                 310                 315                 320
Gly Pro Ile Trp His Gly Val Gln Gly Asn Gly Tyr Ala Leu Ala Asn
                325                 330                 335
Ile Phe Thr Leu Ala Tyr Pro Tyr Gly Tyr Pro Lys Ile Met Ser Gly
            340                 345                 350
Tyr Phe Phe His Gly Asp Phe Asn Ala Ala Pro Pro Ser Ser Gly Ile
        355                 360                 365
His Thr Gly Asn Ala Cys Gly Phe Asp Gly Gly Asp Trp Val Cys Glu
    370                 375                 380
His Lys Trp Arg Gly Ile Ala Asn Met Val Ala Phe Arg Asn Tyr Thr
385                 390                 395                 400
Ala Ser Glu Trp Arg Ile Ser Asn Trp Trp Gln Asn Ser Asn Asp Gln
                405                 410                 415
Ile Ala Phe Gly Arg Gly Gly Leu Gly Phe Val Val Ile Asn Lys Arg
            420                 425                 430
Ala Asn Gly Ser Ile Asn Gln Ser Phe Asp Thr Gly Met Pro Asp Gly
        435                 440                 445
Gln Tyr Cys Asn Ile Ile Glu Ala Asn Phe Asp Glu Ser Thr Gly Gln
    450                 455                 460
Cys Ser Ala Ala Thr Asp Ser Asn Gly Gln Ala Val Ile Thr Val Ser
465                 470                 475                 480
Gly Gly Gln Ala Asn Phe Asn Val Ala Gly Asp His Ala Ala Ala Ile
                485                 490                 495
His Val Gly Ala Lys Ile Gly Asp Gln Cys Ser Gly Asp Asp Cys Pro
            500                 505                 510
Cys Thr Gly Ser Asp Cys Asn Asn Asp Pro Lys Pro Asp Phe Ala Val
        515                 520                 525
Pro Ala Thr Ser Ile Cys Thr Ser Glu Asn Leu Pro Thr Leu Tyr Tyr
    530                 535                 540
Trp Gly Ala Gln Pro Thr Asp Ser Leu Ala Asn Ala Ala Trp Pro Gly
545                 550                 555                 560
Val Ala Met Gln Thr Asn Gly Asp Phe Lys Cys His Asp Leu Gly Val
                565                 570                 575
Glu Leu Thr Lys Ile Asn Ala Ile Phe Ser Asp Asn Gly Ala Asn Lys
            580                 585                 590
Thr Ala Asp Leu Thr Val Thr Gly Ala Gly Cys Tyr Lys Asp Gly Thr
        595                 600                 605
```

Trp Ser Thr Leu Gln Asn Cys Gly Phe Glu Ile Thr Gly Ala Gln Thr
610 615 620
Asn Pro Val Gly Gly Asp Glu Val Trp Tyr Phe Arg Gly Thr Ala Asn
625 630 635 640
Asp Trp Gly Lys Ala Gln Leu Asp Tyr Asp Ala Thr Ser Gly Leu Tyr
645 650 655
Tyr Thr Ile Gln Ser Phe Asn Gly Glu Glu Ala Pro Ala Arg Phe Lys
660 665 670
Ile Asp Asn Gly Ser Trp Thr Glu Ala Tyr Pro Thr Ala Asp Tyr Gln
675 680 685
Val Thr Asp Asn Asn Ser Tyr Arg Ile Asn Phe Asn Ser Asp Ser Lys
690 695 700
Ala Ile Thr Val Asn Ala Gln
705 710

<210> 103
<211> 1650
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 103
gtgctaacgt ttcaccgcat cattcgaaaa ggatggatgt tcctgctcgc gtttttgctc 60
actgcctcgc tgttctgccc aacaggacag cccgccaagg ctgccgcacc gtttaacggc 120
accatgatgc agtattttga atggtacttg ccggatgatg gcacgttatg gaccaaagtg 180
gccaatgaag ccaacaactt atccagcctt ggcatcaccg ctctttggct gccgcccgct 240
tacaaaggaa caagccgcag cgacgtaggg tacggagtat acgacttgta tgacctcggc 300
gaattcaatc aaaaagggac cgtccgcaca aaatacggaa caaaagctca atatcttcaa 360
gccattcaag ccgcccacgc cgctggaatg caagtgtacg ccgatgtcgt gttcgaccat 420
aaaggcggcg ccgacggcac ggaatgggtg gacgccgtcg aagtcaatcc gtccgaccgc 480
aaccaagaaa tctcgggcac ctatcaaatc caagcatgga cgaaatttga ttttcccggg 540
cggggcaaca cctactccag ctttaagtgg cgctggtacc attttgacgg cgttgattgg 600
gacgaaagcc gaaaattgag ccgcatttac aaattccgcg gcatcggcaa agcgtgggat 660
tgggaagtag acacggaaaa cggaaactat gactacttaa tgtatgccga ccttgatatg 720
gatcatcccg aagtcgtgac cgagctgaaa aactgggggg aatggtatgt caacacaacg 780
aacattgatg ggttccggct tgatgccgtc aagcatatta agttcagttt ttttcctgat 840
tggttgtcgt atgtgcgttc tcagactggc aagccgctat ttaccgtcgg ggaatattgg 900
agctatgaca tcaacaagtt gcacaattac attacgaaaa caaacggaac gatgtctttg 960
tttgatgccc cgttacacaa caaatttat accgcttcca aatcaggggg cgcatttgat 1020
atgcgcacgt taatgaccaa tactctcatg aaagatcaac cgacattggc cgtcaccttc 1080
gttgataatc atgacaccga acccggccaa gcgctgcagt catgggtcga cccatggttc 1140
aaaccgttgg cttacgcctt tattctaact cggcaggaag gatacccgtg cgtcttttat 1200
ggtgactatt atggcattcc acaatataac attccttcgc tgaaaagcaa aatcgatccg 1260
ctcctcatcg cgcgcaggga ttatgcttac ggaacgcaac atgattatct tgatcactcc 1320
gacatcatcg ggtggacaag ggaaggggtc actgaaaaac caggatccgg gctggccgca 1380
ctgatcaccg atgggccggg aggaagcaaa tggatgtacg ttggcaaaca acacgctgga 1440
aaagtgttct atgaccttac cggcaaccgg agtgacaccg tcaccatcaa cagtgatgga 1500
tgggggggaat tcaaagtcaa tggcggttcg gtttcggttt gggttcctag aaaaacgacc 1560
gtttctacca tcgctcggcc gatcacaacc cgaccgtgga ctggtgaatt cgtccgttgg 1620
accgaaccac ggttggtggc atggccttga 1650

<210> 104
<211> 549
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 104
Val Leu Thr Phe His Arg Ile Ile Arg Lys Gly Trp Met Phe Leu Leu
1 5 10 15
Ala Phe Leu Leu Thr Ala Ser Leu Phe Cys Pro Thr Gly Gln Pro Ala
20 25 30
Lys Ala Ala Ala Pro Phe Asn Gly Thr Met Met Gln Tyr Phe Glu Trp
35 40 45

```
Tyr Leu Pro Asp Asp Gly Thr Leu Trp Thr Lys Val Ala Asn Glu Ala
    50                  55                  60
Asn Asn Leu Ser Ser Leu Gly Ile Thr Ala Leu Trp Leu Pro Pro Ala
65                  70                  75                  80
Tyr Lys Gly Thr Ser Arg Ser Asp Val Gly Tyr Gly Val Tyr Asp Leu
                85                  90                  95
Tyr Asp Leu Gly Glu Phe Asn Gln Lys Gly Thr Val Arg Thr Lys Tyr
            100                 105                 110
Gly Thr Lys Ala Gln Tyr Leu Gln Ala Ile Gln Ala Ala His Ala Ala
        115                 120                 125
Gly Met Gln Val Tyr Ala Asp Val Val Phe Asp His Lys Gly Gly Ala
    130                 135                 140
Asp Gly Thr Glu Trp Val Asp Ala Val Glu Val Asn Pro Ser Asp Arg
145                 150                 155                 160
Asn Gln Glu Ile Ser Gly Thr Tyr Gln Ile Gln Ala Trp Thr Lys Phe
                165                 170                 175
Asp Phe Pro Gly Arg Gly Asn Thr Tyr Ser Ser Phe Lys Trp Arg Trp
            180                 185                 190
Tyr His Phe Asp Gly Val Asp Trp Asp Glu Ser Arg Lys Leu Ser Arg
        195                 200                 205
Ile Tyr Lys Phe Arg Gly Ile Gly Lys Ala Trp Asp Trp Glu Val Asp
    210                 215                 220
Thr Glu Asn Gly Asn Tyr Asp Tyr Leu Met Tyr Ala Asp Leu Asp Met
225                 230                 235                 240
Asp His Pro Glu Val Val Thr Glu Leu Lys Asn Trp Gly Glu Trp Tyr
                245                 250                 255
Val Asn Thr Thr Asn Ile Asp Gly Phe Arg Leu Asp Ala Val Lys His
            260                 265                 270
Ile Lys Phe Ser Phe Phe Pro Asp Trp Leu Ser Tyr Val Arg Ser Gln
        275                 280                 285
Thr Gly Lys Pro Leu Phe Thr Val Gly Glu Tyr Trp Ser Tyr Asp Ile
    290                 295                 300
Asn Lys Leu His Asn Tyr Ile Thr Lys Thr Asn Gly Thr Met Ser Leu
305                 310                 315                 320
Phe Asp Ala Pro Leu His Asn Lys Phe Tyr Thr Ala Ser Lys Ser Gly
                325                 330                 335
Gly Ala Phe Asp Met Arg Thr Leu Met Thr Asn Thr Leu Met Lys Asp
            340                 345                 350
Gln Pro Thr Leu Ala Val Thr Phe Val Asp Asn His Asp Thr Glu Pro
        355                 360                 365
Gly Gln Ala Leu Gln Ser Trp Val Asp Pro Trp Phe Lys Pro Leu Ala
    370                 375                 380
Tyr Ala Phe Ile Leu Thr Arg Gln Glu Gly Tyr Pro Cys Val Phe Tyr
385                 390                 395                 400
Gly Asp Tyr Tyr Gly Ile Pro Gln Tyr Asn Ile Pro Ser Leu Lys Ser
                405                 410                 415
Lys Ile Asp Pro Leu Leu Ile Ala Arg Arg Asp Tyr Ala Tyr Gly Thr
            420                 425                 430
Gln His Asp Tyr Leu Asp His Ser Asp Ile Ile Gly Trp Thr Arg Glu
        435                 440                 445
Gly Val Thr Glu Lys Pro Gly Ser Gly Leu Ala Ala Leu Ile Thr Asp
    450                 455                 460
Gly Pro Gly Gly Ser Lys Trp Met Tyr Val Gly Lys Gln His Ala Gly
465                 470                 475                 480
Lys Val Phe Tyr Asp Leu Thr Gly Asn Arg Ser Asp Thr Val Thr Ile
                485                 490                 495
Asn Ser Asp Gly Trp Gly Glu Phe Lys Val Asn Gly Gly Ser Val Ser
            500                 505                 510
Val Trp Val Pro Arg Lys Thr Thr Val Ser Thr Ile Ala Arg Pro Ile
        515                 520                 525
Thr Thr Arg Pro Trp Thr Gly Glu Phe Val Arg Trp Thr Glu Pro Arg
    530                 535                 540
Leu Val Ala Trp Pro
545

<210> 105
<211> 1650
<212> DNA
```

<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 105
atgtccctat tcaaaaaaat ctttccgtgg attgtatctc tacttctttt gttttcgttt 60
attgctcctt tttccattca aacagaaaaa gtccgcgctg gaagtgttcc agtgaatgga 120
acgatgatgc aatatttcga atggtacctt ccagacgatg gaacactatg gacgaaagta 180
gcaaataacg cccaatcttt agcgaatctt ggcattactg ccctttggct tccccctgcc 240
tataaaggaa caagcagcag tgacgttgga tatggcgttt atgatttata tgacctagga 300
gagtttaatc aaaaaggaac tgtccgaaca aaatacggaa caaaaacaca atatatccaa 360
gcaatccaag cggcgcatac agcaggaatg caagtatatg cagatgtcgt ctttaaccat 420
aaagccggtg cagatgggac agaactagtg gatgcagtag aagtaaaccc ttctgaccgc 480
aatcaagaaa tatcaggaac atatcaaatc caagcgtgga caaaatttga ttttcctggt 540
cgtggaaaca cctattctag ttttaaatgg cgttggtatc atttcgatgg aacggactgg 600
gatgagagta gaaaactaaa tcgtatttac aaattccgcg gcacgggaaa agcatgggat 660
tgggaagtag atacagaaaa tgggaattat gactatctca tgtatgcaga tttggatatg 720
gatcatccag aggttgtatc tgaactaaaa aattggggaa agtggtatgt aaccacaacc 780
aatatcgacg gattccgtct ggatgcagtg aagcatatta aatatagctt tttcccagac 840
tggctatcgt atgtacgaac ccaaacacaa aagcctcttt ttgccgttgg cgaattttgg 900
agctatgaca ttaacaagct acacaactat attacaaaga cgaacggctc tatgtcccta 960
ttcgatgccc cgctgcataa caatttttat atagcatcga aatcaggtgg ctattttgat 1020
atgcgcacat tactcaacaa cacattgatg aaagatcaac caacactatc ggtcacatta 1080
gtagacaatc acgatactga gccagggcaa tctttgcagt cgtgggtcga gccgtggttt 1140
aaaccgttag cttacgcatt tatcttgacc cgccaagaag gttatccgtg catcttttat 1200
ggagattact atggtattcc aaaatacaac attcctgcgc tgaaaagcaa acttgatccg 1260
ctgttaattg ctcgaagaga ttatgcctac ggaacacagc acgactatat tgacaatgca 1320
gatattatcg gctggacgcg ggaaggagta gctgaaaaag caaattcggg acttgctgca 1380
ctcattaccg acggacctgg cggaagcaaa tggatgtatg ttggcaaaca acacgctggc 1440
aaaacgtttt atgatctaac cggcaatcga agtgatacag tgacaatcaa cgctgatgga 1500
tggggagaat ttaaagtcaa tggagggtct gtatccatat gggttccaaa aacatcaacc 1560
acttcccaaa tcacatttac tgtaaataat gccacaaccg tttggggaca aaatgtatac 1620
gttgtcggga atatttcgca gctgggcaac 1650

<210> 106
<211> 550
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 106
Met Ser Leu Phe Lys Lys Ile Phe Pro Trp Ile Val Ser Leu Leu Leu
 1   5   10   15
Leu Phe Ser Phe Ile Ala Pro Phe Ser Ile Gln Thr Glu Lys Val Arg
            20   25   30
Ala Gly Ser Val Pro Val Asn Gly Thr Met Met Gln Tyr Phe Glu Trp
        35   40   45
Tyr Leu Pro Asp Asp Gly Thr Leu Trp Thr Lys Val Ala Asn Asn Ala
    50   55   60
Gln Ser Leu Ala Asn Leu Gly Ile Thr Ala Leu Trp Leu Pro Pro Ala
65   70   75   80
Tyr Lys Gly Thr Ser Ser Ser Asp Val Gly Tyr Gly Val Tyr Asp Leu
                85   90   95
Tyr Asp Leu Gly Glu Phe Asn Gln Lys Gly Thr Val Arg Thr Lys Tyr
            100   105   110
Gly Thr Lys Thr Gln Tyr Ile Gln Ala Ile Gln Ala Ala His Thr Ala
        115   120   125
Gly Met Gln Val Tyr Ala Asp Val Val Phe Asn His Lys Ala Gly Ala
    130   135   140
Asp Gly Thr Glu Leu Val Asp Ala Val Glu Val Asn Pro Ser Asp Arg
145   150   155   160
Asn Gln Glu Ile Ser Gly Thr Tyr Gln Ile Gln Ala Trp Thr Lys Phe
                165   170   175
Asp Phe Pro Gly Arg Gly Asn Thr Tyr Ser Ser Phe Lys Trp Arg Trp
            180   185   190

```
Tyr His Phe Asp Gly Thr Asp Trp Asp Glu Ser Arg Lys Leu Asn Arg
        195                 200                 205
Ile Tyr Lys Phe Arg Gly Thr Gly Lys Ala Trp Asp Trp Glu Val Asp
    210                 215                 220
Thr Glu Asn Gly Asn Tyr Asp Tyr Leu Met Tyr Ala Asp Leu Asp Met
225                 230                 235                 240
Asp His Pro Glu Val Val Ser Glu Leu Lys Asn Trp Gly Lys Trp Tyr
                245                 250                 255
Val Thr Thr Thr Asn Ile Asp Gly Phe Arg Leu Asp Ala Val Lys His
            260                 265                 270
Ile Lys Tyr Ser Phe Phe Pro Asp Trp Leu Ser Tyr Val Arg Thr Gln
        275                 280                 285
Thr Gln Lys Pro Leu Phe Ala Val Gly Glu Phe Trp Ser Tyr Asp Ile
    290                 295                 300
Asn Lys Leu His Asn Tyr Ile Thr Lys Thr Asn Gly Ser Met Ser Leu
305                 310                 315                 320
Phe Asp Ala Pro Leu His Asn Asn Phe Tyr Ile Ala Ser Lys Ser Gly
                325                 330                 335
Gly Tyr Phe Asp Met Arg Thr Leu Leu Asn Asn Thr Leu Met Lys Asp
            340                 345                 350
Gln Pro Thr Leu Ser Val Thr Leu Val Asp Asn His Asp Thr Glu Pro
        355                 360                 365
Gly Gln Ser Leu Gln Ser Trp Val Glu Pro Trp Phe Lys Pro Leu Ala
    370                 375                 380
Tyr Ala Phe Ile Leu Thr Arg Gln Glu Gly Tyr Pro Cys Ile Phe Tyr
385                 390                 395                 400
Gly Asp Tyr Tyr Gly Ile Pro Lys Tyr Asn Ile Pro Ala Leu Lys Ser
                405                 410                 415
Lys Leu Asp Pro Leu Leu Ile Ala Arg Arg Asp Tyr Ala Tyr Gly Thr
            420                 425                 430
Gln His Asp Tyr Ile Asp Asn Ala Asp Ile Ile Gly Trp Thr Arg Glu
        435                 440                 445
Gly Val Ala Glu Lys Ala Asn Ser Gly Leu Ala Ala Leu Ile Thr Asp
    450                 455                 460
Gly Pro Gly Gly Ser Lys Trp Met Tyr Val Gly Lys Gln His Ala Gly
465                 470                 475                 480
Lys Thr Phe Tyr Asp Leu Thr Gly Asn Arg Ser Asp Thr Val Thr Ile
                485                 490                 495
Asn Ala Asp Gly Trp Gly Glu Phe Lys Val Asn Gly Gly Ser Val Ser
            500                 505                 510
Ile Trp Val Pro Lys Thr Ser Thr Thr Ser Gln Ile Thr Phe Thr Val
        515                 520                 525
Asn Asn Ala Thr Thr Val Trp Gly Gln Asn Val Tyr Val Val Gly Asn
    530                 535                 540
Ile Ser Gln Leu Gly Asn
545                 550

<210> 107
<211> 1509
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 107
atggacagcc tcgacgcgcc ggagcagaag ccctgggtga aggatggcag gctctccgcg      60
tacctggata cagggacagg gaccgtggtc gctcccgagg cacctgcgcc cccgccgccc     120
ccggccgagg aagtccggcc cgtggacaag tggaaaaacg atatcatcta tttcgtcctc     180
accgaccgtt tccaggatgg cgacaagacc aacaacatgg acgtggtccc gacggacatg     240
aaaaaatatc atggcggcga catccagggg ctcatcgaca agctcgacta tatcaaggag     300
accggttcga cggccatctg gctcacgccc cctatgaagg ggcagaccca cttcttcgag     360
accgacaatt accatggtta ctggcccatt gacttctatg acacggaccc ccatgtgggc     420
accatgcaga aatttgagga gcttatcgag aaagcccatg agaaagggct gaagatcgtg     480
ctcgatattc ccctgaacca cacggcctgg gagcatccct tctacaagga cgacagcaag     540
aaggactggt tccaccatat aggagatgtg aaggactggg aagatcccta ctgggctgaa     600
aacggctcca tattcggtct tcctgacctg gcgcaggaaa accctgccgt ggaaaagtac     660
ctcatcgacg tggccaagtt ctgggtagac aagggtattg acggcttcag gcttgacgcc     720
```

gtgaagaacg tgcccctcaa cttctgggcg aagtttgacc gggcgattca cgattatgcg 780
ggcaaggact tcctcctcgt cggggaatac tttgacggaa acccggcgaa agtcgcgaac 840
taccagagag aggacatgag ctcactcttc gattacccgc tctactggac cctgaaggac 900
accttcgcca aggacgggag catgcgcaac ctggcggcga agcttgatga gtgcgacagg 960
aattatcccg acccgggcct catgtcggtt ttccttgata accacgacac gccgaggttc 1020
ctcaccgagg ccaacggcaa caaggataag ctcaaactgg ccctcgcctt cgcgatgacc 1080
atcaaccgca tgcctaccat ttattatggc accgaggttg ccatggaagg caactgcgat 1140
atcatgggcg ccgtagataa ccggagggac atgcagtggg acaaggatcc tgacatgttc 1200
aaatacttca agactctcac cactgcccgc aatgagcatg aatccctcag ggaaggaaag 1260
aagctcgaga tgtggcagga tgacaaagtc tacgcgtacg ggaggcagac cccgaaggac 1320
gagtctatcg tggtgcttaa caacggctat gatacgcagg aacgggacat accgctccgc 1380
cccgagagcg gcatcaagaa cggcacggtg ctgaaggatg tcatcaccgg cgaaaccgtg 1440
acggtacaga acggaaaaat ccatgcgaaa tgcggcggca aacaggcgcg gatctacgtg 1500
cccgcgtag 1509

<210> 108
<211> 502
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 108
Met Asp Ser Leu Asp Ala Pro Glu Gln Lys Pro Trp Val Lys Asp Gly
 1   5   10   15
Arg Leu Ser Ala Tyr Leu Asp Thr Gly Thr Gly Thr Val Val Ala Pro
            20   25   30
Glu Ala Pro Ala Pro Pro Pro Pro Pro Ala Glu Glu Val Arg Pro Val
        35   40   45
Asp Lys Trp Lys Asn Asp Ile Ile Tyr Phe Val Leu Thr Asp Arg Phe
    50   55   60
Gln Asp Gly Asp Lys Thr Asn Asn Met Asp Val Val Pro Thr Asp Met
65   70   75   80
Lys Lys Tyr His Gly Gly Asp Ile Gln Gly Leu Ile Asp Lys Leu Asp
                85   90   95
Tyr Ile Lys Glu Thr Gly Ser Thr Ala Ile Trp Leu Thr Pro Pro Met
            100   105   110
Lys Gly Gln Thr His Phe Phe Glu Thr Asp Asn Tyr His Gly Tyr Trp
        115   120   125
Pro Ile Asp Phe Tyr Asp Thr Asp Pro His Val Gly Thr Met Gln Lys
    130   135   140
Phe Glu Glu Leu Ile Glu Lys Ala His Glu Lys Gly Leu Lys Ile Val
145   150   155   160
Leu Asp Ile Pro Leu Asn His Thr Ala Trp Glu His Pro Phe Tyr Lys
                165   170   175
Asp Asp Ser Lys Lys Asp Trp Phe His His Ile Gly Asp Val Lys Asp
            180   185   190
Trp Glu Asp Pro Tyr Trp Ala Glu Asn Gly Ser Ile Phe Gly Leu Pro
        195   200   205
Asp Leu Ala Gln Glu Asn Pro Ala Val Glu Lys Tyr Leu Ile Asp Val
    210   215   220
Ala Lys Phe Trp Val Asp Lys Gly Ile Asp Gly Phe Arg Leu Asp Ala
225   230   235   240
Val Lys Asn Val Pro Leu Asn Phe Trp Ala Lys Phe Asp Arg Ala Ile
                245   250   255
His Asp Tyr Ala Gly Lys Asp Phe Leu Leu Val Gly Glu Tyr Phe Asp
            260   265   270
Gly Asn Pro Ala Lys Val Ala Asn Tyr Gln Arg Glu Asp Met Ser Ser
        275   280   285
Leu Phe Asp Tyr Pro Leu Tyr Trp Thr Leu Lys Asp Thr Phe Ala Lys
    290   295   300
Asp Gly Ser Met Arg Asn Leu Ala Ala Lys Leu Asp Glu Cys Asp Arg
305   310   315   320
Asn Tyr Pro Asp Pro Gly Leu Met Ser Val Phe Leu Asp Asn His Asp
                325   330   335
Thr Pro Arg Phe Leu Thr Glu Ala Asn Gly Asn Lys Asp Lys Leu Lys
            340   345   350

```
Leu Ala Leu Ala Phe Ala Met Thr Ile Asn Arg Met Pro Thr Ile Tyr
        355                 360                 365
Tyr Gly Thr Glu Val Ala Met Glu Gly Asn Cys Asp Ile Met Gly Ala
    370                 375                 380
Val Asp Asn Arg Arg Asp Met Gln Trp Asp Lys Asp Pro Asp Met Phe
385                 390                 395                 400
Lys Tyr Phe Lys Thr Leu Thr Thr Ala Arg Asn Glu His Glu Ser Leu
                405                 410                 415
Arg Glu Gly Lys Lys Leu Glu Met Trp Gln Asp Asp Lys Val Tyr Ala
            420                 425                 430
Tyr Gly Arg Gln Thr Pro Lys Asp Glu Ser Ile Val Val Leu Asn Asn
        435                 440                 445
Gly Tyr Asp Thr Gln Glu Arg Asp Ile Pro Leu Arg Pro Glu Ser Gly
    450                 455                 460
Ile Lys Asn Gly Thr Val Leu Lys Asp Val Ile Thr Gly Glu Thr Val
465                 470                 475                 480
Thr Val Gln Asn Gly Lys Ile His Ala Lys Cys Gly Gly Lys Gln Ala
                485                 490                 495
Arg Ile Tyr Val Pro Ala
            500
```

<210> 109
<211> 1374
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 109

```
atggcaagaa aaacgctggc catatttttc gtacttctag tgcttcttag tctctcggca      60
gttccggcaa aggcagaaac tctagagaat ggtggagtta taatgcaggc tttctattgg     120
gatgttcctg gaggaggaat ctggtgggac acaatagctc aaaagatacc cgaatgggca     180
agtgcaggaa tctcagcgat atggattcca ccagcgagta agggcatgag cggtggttat     240
tccatgggct acgatcccta cgatttcttt gacctcggcg agtactatca gaaggggaca     300
gttgagacgc gcttcggctc aaaggaagaa ctggtgaaca tgataaacac cgcacactcc     360
tacggcataa aggtgatagc ggacatagtc ataaaccacc gcgccggtgg agaccttgag     420
tggaacccct tcgtgaacga ctatacctgg acagacttct caaaagtcgc ctccggtaaa     480
tatacggcca actaccttga cttccaccca aacgagcttc actgttgtga tgaaggtacc     540
tttggaggat accctgatat atgtcacgac aaaagctggg accagtactg gctctgggcg     600
agcagcgaaa gctacgctgc ctacctcagg agcatagggg ttgacgcctg gcgtttcgac     660
tacgtcaagg gctacggagc atgggttgtt aacgactggc tcagctggtg gggaggctgg     720
gccgttggag agtactggga cacgaacgtt gatgcactcc tcaactgggc atacagcagc     780
ggcgccaagg tctttgactt cccgctctac tacaagatgg acgaagcctt cgacaacacc     840
aacatcccgg cattagtgga tgcactcaga tacggccaga cagtggtcag ccgcgatccc     900
ttcaaggcgg taactttcgt tgccaaccac gatacagata taatctggaa caagtatccg     960
gcttatgcat tcatccttac ctatgaggga cagcctgtta tattctaccg cgactacgag    1020
gagtggctca acaaggataa gcttaacaac ctcatctgga tacacgatca ccttgctgga    1080
gggagtactg acattgttta ctacgacagc gacgagctta tctttgtgag aaacggctat    1140
ggcaccaaac caggactgat aacctatatc aacctcggct caagcaaagt tggaaggtgg    1200
gtctacgttc caaagttcgc cggttcatgc atccacgagt acaccggcaa cctcggcggt    1260
tggatagaca agtacgtctc ctccagcggc tgggtctatc ttgaggcccc agcccacgac    1320
ccggcgaacg gctactacgg ctactctgtc tggagctact gcggtgtggg ttga          1374
```

<210> 110
<211> 457
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 110

```
Met Ala Arg Lys Thr Leu Ala Ile Phe Phe Val Leu Leu Val Leu Leu
 1               5                  10                  15
Ser Leu Ser Ala Val Pro Ala Lys Ala Glu Thr Leu Glu Asn Gly Gly
            20                  25                  30
Val Ile Met Gln Ala Phe Tyr Trp Asp Val Pro Gly Gly Gly Ile Trp
```

```
        35                  40                  45
Trp Asp Thr Ile Ala Gln Lys Ile Pro Glu Trp Ala Ser Ala Gly Ile
    50                  55                  60
Ser Ala Ile Trp Ile Pro Pro Ala Ser Lys Gly Met Ser Gly Gly Tyr
65                  70                  75                  80
Ser Met Gly Tyr Asp Pro Tyr Asp Phe Phe Asp Leu Gly Glu Tyr Tyr
                85                  90                  95
Gln Lys Gly Thr Val Glu Thr Arg Phe Gly Ser Lys Glu Glu Leu Val
            100                 105                 110
Asn Met Ile Asn Thr Ala His Ser Tyr Gly Ile Lys Val Ile Ala Asp
        115                 120                 125
Ile Val Ile Asn His Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe
    130                 135                 140
Val Asn Asp Tyr Thr Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys
145                 150                 155                 160
Tyr Thr Ala Asn Tyr Leu Asp Phe His Pro Asn Glu Leu His Cys Cys
                165                 170                 175
Asp Glu Gly Thr Phe Gly Gly Tyr Pro Asp Ile Cys His Asp Lys Ser
            180                 185                 190
Trp Asp Gln Tyr Trp Leu Trp Ala Ser Ser Glu Ser Tyr Ala Ala Tyr
        195                 200                 205
Leu Arg Ser Ile Gly Val Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly
    210                 215                 220
Tyr Gly Ala Trp Val Val Asn Asp Trp Leu Ser Trp Trp Gly Gly Trp
225                 230                 235                 240
Ala Val Gly Glu Tyr Trp Asp Thr Asn Val Asp Ala Leu Leu Asn Trp
                245                 250                 255
Ala Tyr Ser Ser Gly Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys
            260                 265                 270
Met Asp Glu Ala Phe Asp Asn Thr Asn Ile Pro Ala Leu Val Asp Ala
        275                 280                 285
Leu Arg Tyr Gly Gln Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val
    290                 295                 300
Thr Phe Val Ala Asn His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro
305                 310                 315                 320
Ala Tyr Ala Phe Ile Leu Thr Tyr Glu Gly Gln Pro Val Ile Phe Tyr
                325                 330                 335
Arg Asp Tyr Glu Glu Trp Leu Asn Lys Asp Lys Leu Asn Asn Leu Ile
            340                 345                 350
Trp Ile His Asp His Leu Ala Gly Gly Ser Thr Asp Ile Val Tyr Tyr
        355                 360                 365
Asp Ser Asp Glu Leu Ile Phe Val Arg Asn Gly Tyr Gly Thr Lys Pro
    370                 375                 380
Gly Leu Ile Thr Tyr Ile Asn Leu Gly Ser Ser Lys Val Gly Arg Trp
385                 390                 395                 400
Val Tyr Val Pro Lys Phe Ala Gly Ser Cys Ile His Glu Tyr Thr Gly
                405                 410                 415
Asn Leu Gly Gly Trp Ile Asp Lys Tyr Val Ser Ser Ser Gly Trp Val
            420                 425                 430
Tyr Leu Glu Ala Pro Ala His Asp Pro Ala Asn Gly Tyr Tyr Gly Tyr
        435                 440                 445
Ser Val Trp Ser Tyr Cys Gly Val Gly
    450                 455

<210> 111
<211> 1416
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 111
atgcccgcgt tcaaatctaa ggtgatgcac atgaagttga agtaccttgc tttagttttg      60
ttggctgtgg cttcgatagg cctcctctcg actccagtgg gtgctgccaa gtactccgaa     120
ctcgaagagg gcggtgttat aatgcaggcc ttctactggg acgtccctac cggtgggatc     180
tggtgggaca ccataagaca gaaaatcccg gagtggtacg acgctggaat ctcggcgata     240
tggattcctc cagctagcaa aggtatgggt ggtgcatact ccatgggtta tgacccctac     300
```

```
gatttctttg acctcggcga gtactatcag aagggaacag ttgagacgcg cttcggctca     360
aaggaggaac tggtgaacat gataaacacc gcacactcct atggcataaa ggtgatagcg     420
gacatagtca taaaccaccg cgccggcggc gacctggagt ggaacccctt tgtaaacaac     480
tatacttgga cagacttctc caaggtcgcc tccggtaaat acacggccaa ctaccttgac     540
ttccacccaa acgaggtcaa gtgctgcgat gagggtacat ttggtgactt tccggacatc     600
gcccacgaga agagctggga tcagtactgg ctctgggcaa gcaatgagag ctacgccgcc     660
tatctccgga gcatagggat cgatgcatgg cgtttcgact acgtcaaagg ttacggagcg     720
tgggttgtta acgactggct cagctggtgg ggaggttggg ccgttggaga gtactgggac     780
accaacgttg atgcactcct taactgggca tacaacagcg gtgccaaggt ctttgacttc     840
ccgctctact acaagatgga cgaagccttt gacaacacca acatccccgc tttggtttac     900
gccctccaga acggaggaac agtcgtttcc cgcgatccct tcaaggcagt aactttcgtt     960
gccaaccacg ataccgatat aatctggaac aagtatccgg cttatgcgtt catccttacc    1020
tatgagggac agcctgttat attctaccgc gactacgagg agtggctcaa caaggataag    1080
cttaacaacc ttatctggat acacgagcac cttgccggag gaagtaccaa gatcctctac    1140
tacgataacg atgagctaat attcatgagg gagggctacg ggagcaagcc gggcctcata    1200
acctacataa acctcggaaa cgactgggcc gagcgctggg tgaacgtcgg ctcaaagttt    1260
gccggctaca caatccatga atacacaggc aatctcggtg gctgggttga caggtgggtt    1320
cagtacgacg gatgggttaa actgacggca cctcctcacg atccagccaa cggatattac    1380
ggctactcag tctggagcta cgcaggcgtc ggatga                              1416

<210> 112
<211> 471
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 112
Met Pro Ala Phe Lys Ser Lys Val Met His Met Lys Leu Lys Tyr Leu
 1               5                  10                  15
Ala Leu Val Leu Leu Ala Val Ala Ser Ile Gly Leu Leu Ser Thr Pro
            20                  25                  30
Val Gly Ala Ala Lys Tyr Ser Glu Leu Glu Glu Gly Gly Val Ile Met
        35                  40                  45
Gln Ala Phe Tyr Trp Asp Val Pro Thr Gly Gly Ile Trp Trp Asp Thr
    50                  55                  60
Ile Arg Gln Lys Ile Pro Glu Trp Tyr Asp Ala Gly Ile Ser Ala Ile
65                  70                  75                  80
Trp Ile Pro Pro Ala Ser Lys Gly Met Gly Gly Ala Tyr Ser Met Gly
                85                  90                  95
Tyr Asp Pro Tyr Asp Phe Phe Asp Leu Gly Glu Tyr Tyr Gln Lys Gly
            100                 105                 110
Thr Val Glu Thr Arg Phe Gly Ser Lys Glu Glu Leu Val Asn Met Ile
        115                 120                 125
Asn Thr Ala His Ser Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile
    130                 135                 140
Asn His Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Asn Asn
145                 150                 155                 160
Tyr Thr Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala
                165                 170                 175
Asn Tyr Leu Asp Phe His Pro Asn Glu Val Lys Cys Cys Asp Glu Gly
            180                 185                 190
Thr Phe Gly Asp Phe Pro Asp Ile Ala His Glu Lys Ser Trp Asp Gln
        195                 200                 205
Tyr Trp Leu Trp Ala Ser Asn Glu Ser Tyr Ala Ala Tyr Leu Arg Ser
    210                 215                 220
Ile Gly Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala
225                 230                 235                 240
Trp Val Val Asn Asp Trp Leu Ser Trp Trp Gly Gly Trp Ala Val Gly
                245                 250                 255
Glu Tyr Trp Asp Thr Asn Val Asp Ala Leu Leu Asn Trp Ala Tyr Asn
            260                 265                 270
Ser Gly Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Glu
        275                 280                 285
Ala Phe Asp Asn Thr Asn Ile Pro Ala Leu Val Tyr Ala Leu Gln Asn
    290                 295                 300
Gly Gly Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val
```

305 310 315 320
Ala Asn His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala
325 330 335
Phe Ile Leu Thr Tyr Glu Gly Gln Pro Val Ile Phe Tyr Arg Asp Tyr
340 345 350
Glu Glu Trp Leu Asn Lys Asp Lys Leu Asn Asn Leu Ile Trp Ile His
355 360 365
Glu His Leu Ala Gly Gly Ser Thr Lys Ile Leu Tyr Tyr Asp Asn Asp
370 375 380
Glu Leu Ile Phe Met Arg Glu Gly Tyr Gly Ser Lys Pro Gly Leu Ile
385 390 395 400
Thr Tyr Ile Asn Leu Gly Asn Asp Trp Ala Glu Arg Trp Val Asn Val
405 410 415
Gly Ser Lys Phe Ala Gly Tyr Thr Ile His Glu Tyr Thr Gly Asn Leu
420 425 430
Gly Gly Trp Val Asp Arg Trp Val Gln Tyr Asp Gly Trp Val Lys Leu
435 440 445
Thr Ala Pro Pro His Asp Pro Ala Asn Gly Tyr Tyr Gly Tyr Ser Val
450 455 460
Trp Ser Tyr Ala Gly Val Gly
465 470

<210> 113
<211> 1539
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 113
atgaaacaac aaaaacggct ttacgcccga ttgctgacgc tgttatttgc gctcatcttc 60
ttgctgcctc attctgcagc agcggcggca aatcttaatg ggacgctgat gcagtatttt 120
gaatggtaca tgcccaatga cggccaacat tggaagcgct tgcaaaacga ctcggcatat 180
ttggctgaac acggtattac tgccgtctgg attcccccgg catataaggg aacgagccaa 240
gcggatgtgg gctacggtgc ttacgacctt tatgatttag gggagtttca tcaaaaaggg 300
acggttcgga caaagtacgg cacaaaagga gagctgcaat ctgcgatcaa aagtcttcat 360
tcccgcgaca ttaacgttta cggggatgtg gtcatcaacc acaaaggcgg cgctgatgcg 420
accgaagatg taaccgcggt tgaagtcgat cccgctgacc gcaaccgcgt aatttcagga 480
gaacaccgaa ttaaagcctg gacacatttt cattttccgg ggcgcggcag cacatacagc 540
gattttaaat ggcattggta ccattttgac ggaaccgatt gggacgagtc ccgaaagctg 600
aaccgcatct ataagtttca aggaaaggct tgggattggg aagtttccaa tgaaaacggc 660
aactatgatt atttgatgta tgccgacatc gattatgacc atcctgatgt cgcagcagaa 720
attaagagat ggggcacttg gtatgccaat gaactgcaat tggacggttt ccgtcttgat 780
gctgtcaaac acattaaatt ttctttttg cgggattggg ttaatcatgt cagggaaaaa 840
acggggaagg aaatgtttac ggtagctgaa tattggcaga atgacttggg cgcgctggaa 900
aactatttga acaaaacaaa ttttaatcat tcagtgtttg acgtgccgct tcattatcag 960
ttccatgctg catcgacaca gggaggcggc tatgatatga ggaaattgct gaacggtacg 1020
gtcgtttcca agcatccgtt gaaagcggtt acatttgtcg ataaccatga tacacagccg 1080
gggcaatcgc ttgagtcgac tgtccaaaca tggtttaagc cgcttgctta cgctttcatt 1140
ctcacaaggg aatctggata ccctcaggtt ttctacgggg atatgtacgg gacgaaagga 1200
gactcccagc gcgaaattcc tgccttgaaa cacaaaattg aaccgatctt aaaagcgaga 1260
aaacagtatg cgtacggagc acagcatgat tatttcgacc accatgacat tgtcggctgg 1320
acaagggaag gcgacagctc ggttgcaaat tcaggtttgg cggcattaat aacagacgga 1380
cccggtgggg caaagcgaat gtatgtcggc cggcaaaacg ccggtgagac atggcatgac 1440
attaccggaa accgttcgga gccggttgtc atcaattcgg aaggctgggg agagtttcac 1500
gtaaacggcg ggtcggtttc aatttatgtt caaagatag 1539

<210> 114
<211> 512
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 114
Met Lys Gln Gln Lys Arg Leu Tyr Ala Arg Leu Leu Thr Leu Leu Phe

```
 1               5                  10                  15
Ala Leu Ile Phe Leu Leu Pro His Ser Ala Ala Ala Ala Ala Asn Leu
            20                  25                  30
Asn Gly Thr Leu Met Gln Tyr Phe Glu Trp Tyr Met Pro Asn Asp Gly
        35                  40                  45
Gln His Trp Lys Arg Leu Gln Asn Asp Ser Ala Tyr Leu Ala Glu His
    50                  55                  60
Gly Ile Thr Ala Val Trp Ile Pro Pro Ala Tyr Lys Gly Thr Ser Gln
65                  70                  75                  80
Ala Asp Val Gly Tyr Gly Ala Tyr Asp Leu Tyr Asp Leu Gly Glu Phe
                85                  90                  95
His Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly Thr Lys Gly Glu Leu
            100                 105                 110
Gln Ser Ala Ile Lys Ser Leu His Ser Arg Asp Ile Asn Val Tyr Gly
        115                 120                 125
Asp Val Val Ile Asn His Lys Gly Gly Ala Asp Ala Thr Glu Asp Val
    130                 135                 140
Thr Ala Val Glu Val Asp Pro Ala Asp Arg Asn Arg Val Ile Ser Gly
145                 150                 155                 160
Glu His Arg Ile Lys Ala Trp Thr His Phe His Phe Pro Gly Arg Gly
                165                 170                 175
Ser Thr Tyr Ser Asp Phe Lys Trp His Trp Tyr His Phe Asp Gly Thr
            180                 185                 190
Asp Trp Asp Glu Ser Arg Lys Leu Asn Arg Ile Tyr Lys Phe Gln Gly
        195                 200                 205
Lys Ala Trp Asp Trp Glu Val Ser Asn Glu Asn Gly Asn Tyr Asp Tyr
    210                 215                 220
Leu Met Tyr Ala Asp Ile Asp Tyr Asp His Pro Asp Val Ala Ala Glu
225                 230                 235                 240
Ile Lys Arg Trp Gly Thr Trp Tyr Ala Asn Glu Leu Gln Leu Asp Gly
                245                 250                 255
Phe Arg Leu Asp Ala Val Lys His Ile Lys Phe Ser Phe Leu Arg Asp
            260                 265                 270
Trp Val Asn His Val Arg Glu Lys Thr Gly Lys Glu Met Phe Thr Val
        275                 280                 285
Ala Glu Tyr Trp Gln Asn Asp Leu Gly Ala Leu Glu Asn Tyr Leu Asn
    290                 295                 300
Lys Thr Asn Phe Asn His Ser Val Phe Asp Val Pro Leu His Tyr Gln
305                 310                 315                 320
Phe His Ala Ala Ser Thr Gln Gly Gly Gly Tyr Asp Met Arg Lys Leu
                325                 330                 335
Leu Asn Gly Thr Val Val Ser Lys His Pro Leu Lys Ala Val Thr Phe
            340                 345                 350
Val Asp Asn His Asp Thr Gln Pro Gly Gln Ser Leu Glu Ser Thr Val
        355                 360                 365
Gln Thr Trp Phe Lys Pro Leu Ala Tyr Ala Phe Ile Leu Thr Arg Glu
    370                 375                 380
Ser Gly Tyr Pro Gln Val Phe Tyr Gly Asp Met Tyr Gly Thr Lys Gly
385                 390                 395                 400
Asp Ser Gln Arg Glu Ile Pro Ala Leu Lys His Lys Ile Glu Pro Ile
                405                 410                 415
Leu Lys Ala Arg Lys Gln Tyr Ala Tyr Gly Ala Gln His Asp Tyr Phe
            420                 425                 430
Asp His His Asp Ile Val Gly Trp Thr Arg Glu Gly Asp Ser Ser Val
        435                 440                 445
Ala Asn Ser Gly Leu Ala Ala Leu Ile Thr Asp Gly Pro Gly Gly Ala
    450                 455                 460
Lys Arg Met Tyr Val Gly Arg Gln Asn Ala Gly Glu Thr Trp His Asp
465                 470                 475                 480
Ile Thr Gly Asn Arg Ser Glu Pro Val Val Ile Asn Ser Glu Gly Trp
                485                 490                 495
Gly Glu Phe His Val Asn Gly Gly Ser Val Ser Ile Tyr Val Gln Arg
            500                 505                 510
```

<210> 115
<211> 1338
<212> DNA
<213> Eukaryote

<400> 115
atggcgaagt actccgagct ggagcagggc ggagtcataa tgcaggcctt ctactgggac 60
gttccggagg gaggaatctg gtgggacaca atacggcaga agatccctga atggtacgat 120
gcaggcatat ccgccatctg gataccccg gcgagcaagg gcatgggcgg ggcctactcg 180
atgggctacg acccctacga ttacttcgat ctgggcgagt tttaccagaa gggaaccgtt 240
gagacccgct tcggctccaa ggaagagctc gtcaacatga tctccacggc ccaccagtac 300
ggcatcaagg ttatagcgga catagtgata aaccaccgcg caggtggaga cctcgaatgg 360
aacccatacg tcggcgacta tacctggacg gacttttcta aggtcgcctc cgggaaatac 420
aaggcccact acatggactt ccatccaaac aactacagca cctcagacga gggaaccttc 480
ggtggcttcc cagacattga tcacctcgtg cccttcaacc agtactggct gtgggcgagc 540
aacgagagct acgccgccta cctcaggagc atagggatcg atgcgtggcg ctttgactac 600
gttaagggct acggcgcgtg ggtcgtcaag gactggctga gtcagtgggg cggctgggcc 660
gtcggcgagt actgggacac caacgtcgat gcgctcctca actgggccta cagcagcggc 720
gccaaggtct tcgacttccc gctctactac aagatggacg aggcctttga caacaagaac 780
attcccgccc tcgtttacgc catccagaac ggtgaaaccg tcgtcagcag ggatcccttc 840
aaggccgtta ccttcgtggc taaccacgat acgaacataa tctggaacaa gtaccctgcc 900
tatgccttca tcctgaccta cgaaggtcag cccgtcatct tctaccgcga ctacgaggag 960
tggctcaaca aggacaaact caacaacctc atatggattc acgagcacct ggcaggggga 1020
agcaccaaga tcctctacta cgacgacgat gagctcatct tcatgaggga aggctacggc 1080
gacaggcccg ggcttataac ctacatcaac ctcggtagcg actgggcgga gagatgggtg 1140
aacgttggct caaagttcgc gggctataca atccacgaat acaccggaaa cctcggcggc 1200
tgggtcgaca ggtacgtcca gtacgacggc tgggtcaagc ttaccgctcc gccacacgat 1260
ccggcaaacg gctattacgg ctactcggtc tggagctacg ccggagttgg aagatctcat 1320
caccatcacc atcactaa 1338

<210> 116
<211> 445
<212> PRT
<213> Eukaryote

<400> 116
Met Ala Lys Tyr Ser Glu Leu Glu Gln Gly Gly Val Ile Met Gln Ala
 1   5   10   15
Phe Tyr Trp Asp Val Pro Glu Gly Gly Ile Trp Trp Asp Thr Ile Arg
            20   25   30
Gln Lys Ile Pro Glu Trp Tyr Asp Ala Gly Ile Ser Ala Ile Trp Ile
        35   40   45
Pro Pro Ala Ser Lys Gly Met Gly Gly Ala Tyr Ser Met Gly Tyr Asp
    50   55   60
Pro Tyr Asp Tyr Phe Asp Leu Gly Glu Phe Tyr Gln Lys Gly Thr Val
65   70   75   80
Glu Thr Arg Phe Gly Ser Lys Glu Glu Leu Val Asn Met Ile Ser Thr
                85   90   95
Ala His Gln Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His
            100   105   110
Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Tyr Val Gly Asp Tyr Thr
        115   120   125
Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Lys Ala His Tyr
    130   135   140
Met Asp Phe His Pro Asn Asn Tyr Ser Thr Ser Asp Glu Gly Thr Phe
145   150   155   160
Gly Gly Phe Pro Asp Ile Asp His Leu Val Pro Phe Asn Gln Tyr Trp
                165   170   175
Leu Trp Ala Ser Asn Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly
            180   185   190
Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp Val
        195   200   205
Val Lys Asp Trp Leu Ser Gln Trp Gly Gly Trp Ala Val Gly Glu Tyr
    210   215   220
Trp Asp Thr Asn Val Asp Ala Leu Leu Asn Trp Ala Tyr Ser Ser Gly
225   230   235   240
Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Glu Ala Phe
                245   250   255
Asp Asn Lys Asn Ile Pro Ala Leu Val Tyr Ala Ile Gln Asn Gly Glu
            260   265   270
Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn
        275   280   285

```
His Asp Thr Asn Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile
    290                 295                 300
Leu Thr Tyr Glu Gly Gln Pro Val Ile Phe Tyr Arg Asp Tyr Glu Glu
305                 310                 315                 320
Trp Leu Asn Lys Asp Lys Leu Asn Asn Leu Ile Trp Ile His Glu His
                325                 330                 335
Leu Ala Gly Gly Ser Thr Lys Ile Leu Tyr Tyr Asp Asp Asp Glu Leu
            340                 345                 350
Ile Phe Met Arg Glu Gly Tyr Gly Asp Arg Pro Gly Leu Ile Thr Tyr
        355                 360                 365
Ile Asn Leu Gly Ser Asp Trp Ala Glu Arg Trp Val Asn Val Gly Ser
    370                 375                 380
Lys Phe Ala Gly Tyr Thr Ile His Glu Tyr Thr Gly Asn Leu Gly Gly
385                 390                 395                 400
Trp Val Asp Arg Tyr Val Gln Tyr Asp Gly Trp Val Lys Leu Thr Ala
                405                 410                 415
Pro Pro His Asp Pro Ala Asn Gly Tyr Tyr Gly Tyr Ser Val Trp Ser
            420                 425                 430
Tyr Ala Gly Val Gly Arg Ser His His His His His His
        435                 440                 445
```

<210> 117
<211> 1476
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 117

```
atgcgagtgt tcctggttgt gccaaagctg agccgcccat ttcaggcaga gtcacaacaa      60
caagacaggg acataacaat gaaacacaca gcgggaatgc tggcgatcgc aggtatgctg     120
atcgcccccc tggcgcatgc cgatgtcata ctgcacgcct tcaactggaa atacagtgaa     180
gtcaccgcca aggccgatct catcaaggct gccggctaca agcaggtgct catctcaccg     240
cctctgaagt cctcgggcaa cgagtggtgg gctcgttacc agccccagga tctgcgcctg     300
gtcgacaccc cccttggcaa caagcaggat ctggagcagc tgatcgccgc gatgcagacc     360
cggggcattg ccgtctacgc ggacgtggtg ctcaaccaca tggccaacga aagctggaag     420
cgcagcgacc tcaactaccc cggcagcgag ctgctgcaaa gctacgccgg caatccggcc     480
tactttgaac gccagaagct ctttggcgat ctggggcaga acttcctcgc cggccaggat     540
tttcatccgg aggggtgcat caccgactgg aacaatccgg gccatgtcca gtactggcga     600
ctgtgcggcg gggcgggtga caaggggctg ccggatctgg accccaacaa ctgggtggtg     660
aaccagcaac aggcttacct gcaggcgctc aaggggatgg ggatcaaggg ttttcgggtc     720
gatgcggtca agcacatgag cgattaccag atcaacgccg tgttcacccc cgagatcaaa     780
caggggatgc acgtctttgg cgaggtgatc accacggggg gcgccggcaa cagcgactat     840
gagaacttcc tcaaacccta cctcgacagc agcggccagg gggcctacga cttcccgctc     900
ttcgcctccc tgcgtggagc gctgggctac ggcggcagca tgaacctgct ggccgatccc     960
ggtgcctatg gtcaggcgct gccgggtagc cgcgccgtca ccttcgccat cacccacgac    1020
atccccacca acgacggttt ccgctaccag atcctcaacc agaccgacga gagactggcc    1080
tatgcctacc tgctcggtcg cgatggcggt tcgcctctgg tctactccga tcacggtgaa    1140
accagggaca aggacggatt gcgctggcag gactactatc tgcgcaccga tctcaaaggg    1200
atgatccgct tccataacac agtgcagggt caaccgatgc agctcatcgg cagtaacgac    1260
tgcttcgtgc tgttcaagcg tggcaagcag ggcgtggtcg gcatcaacaa gtgcgactac    1320
gagcaggagt actggctcga taccgccaga ttcgagatga actggtatcg caactaccgg    1380
gatgtgctcg accagaatgc cgtggtcaac gtgcagagcc agtgggtaag gctgaccatc    1440
ccggcccgcg gcgccagaat gtggctgcag gagtga                               1476
```

<210> 118
<211> 491
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 118

```
Met Arg Val Phe Leu Val Val Pro Lys Leu Ser Arg Pro Phe Gln Ala
 1               5                  10                  15
Glu Ser Gln Gln Gln Asp Arg Asp Ile Thr Met Lys His Thr Ala Gly
```

```
            20                  25                  30
Met Leu Ala Ile Ala Gly Met Leu Ile Ala Pro Leu Ala His Ala Asp
        35                  40                  45
Val Ile Leu His Ala Phe Asn Trp Lys Tyr Ser Glu Val Thr Ala Lys
    50                  55                  60
Ala Asp Leu Ile Lys Ala Ala Gly Tyr Lys Gln Val Leu Ile Ser Pro
65                  70                  75                  80
Pro Leu Lys Ser Ser Gly Asn Glu Trp Trp Ala Arg Tyr Gln Pro Gln
                85                  90                  95
Asp Leu Arg Leu Val Asp Thr Pro Leu Gly Asn Lys Gln Asp Leu Glu
            100                 105                 110
Gln Leu Ile Ala Ala Met Gln Thr Arg Gly Ile Ala Val Tyr Ala Asp
        115                 120                 125
Val Val Leu Asn His Met Ala Asn Glu Ser Trp Lys Arg Ser Asp Leu
    130                 135                 140
Asn Tyr Pro Gly Ser Glu Leu Leu Gln Ser Tyr Ala Gly Asn Pro Ala
145                 150                 155                 160
Tyr Phe Glu Arg Gln Lys Leu Phe Gly Asp Leu Gly Gln Asn Phe Leu
                165                 170                 175
Ala Gly Gln Asp Phe His Pro Glu Gly Cys Ile Thr Asp Trp Asn Asn
            180                 185                 190
Pro Gly His Val Gln Tyr Trp Arg Leu Cys Gly Gly Ala Gly Asp Lys
        195                 200                 205
Gly Leu Pro Asp Leu Asp Pro Asn Asn Trp Val Val Asn Gln Gln Gln
    210                 215                 220
Ala Tyr Leu Gln Ala Leu Lys Gly Met Gly Ile Lys Gly Phe Arg Val
225                 230                 235                 240
Asp Ala Val Lys His Met Ser Asp Tyr Gln Ile Asn Ala Val Phe Thr
                245                 250                 255
Pro Glu Ile Lys Gln Gly Met His Val Phe Gly Glu Val Ile Thr Thr
            260                 265                 270
Gly Gly Ala Gly Asn Ser Asp Tyr Glu Asn Phe Leu Lys Pro Tyr Leu
        275                 280                 285
Asp Ser Ser Gly Gln Gly Ala Tyr Asp Phe Pro Leu Phe Ala Ser Leu
    290                 295                 300
Arg Gly Ala Leu Gly Tyr Gly Gly Ser Met Asn Leu Leu Ala Asp Pro
305                 310                 315                 320
Gly Ala Tyr Gly Gln Ala Leu Pro Gly Ser Arg Ala Val Thr Phe Ala
                325                 330                 335
Ile Thr His Asp Ile Pro Thr Asn Asp Gly Phe Arg Tyr Gln Ile Leu
            340                 345                 350
Asn Gln Thr Asp Glu Arg Leu Ala Tyr Ala Tyr Leu Leu Gly Arg Asp
        355                 360                 365
Gly Gly Ser Pro Leu Val Tyr Ser Asp His Gly Glu Thr Arg Asp Lys
    370                 375                 380
Asp Gly Leu Arg Trp Gln Asp Tyr Tyr Leu Arg Thr Asp Leu Lys Gly
385                 390                 395                 400
Met Ile Arg Phe His Asn Thr Val Gln Gly Gln Pro Met Gln Leu Ile
                405                 410                 415
Gly Ser Asn Asp Cys Phe Val Leu Phe Lys Arg Gly Lys Gln Gly Val
            420                 425                 430
Val Gly Ile Asn Lys Cys Asp Tyr Glu Gln Glu Tyr Trp Leu Asp Thr
        435                 440                 445
Ala Arg Phe Glu Met Asn Trp Tyr Arg Asn Tyr Arg Asp Val Leu Asp
    450                 455                 460
Gln Asn Ala Val Val Asn Val Gln Ser Gln Trp Val Arg Leu Thr Ile
465                 470                 475                 480
Pro Ala Arg Gly Ala Arg Met Trp Leu Gln Glu
                485                 490
```

<210> 119
<211> 1695
<212> DNA
<213> Bacterial

<400> 119
atgcaaacgt ttgcattctt attttactca aagaaaggat gggtgtgcat gaattatttg 60
aaaaaagtgt ggttgtatta cgctatcgtc gctaccttaa tcatttcctt tcttacacct 120

```
ttttcaacag cacaagctaa tactgcacct gttaacggaa caatgatgca atatttcgaa     180
tgggacttac ctaatgatgg gacgctttgg acgaaagtaa aaaatgaagc taccaatctt     240
tcttcactag gtatcacagc actatggctc cctccagcat ataaaggaac gagccaaagc     300
gatgtcggat acggtgttta cgatttatat gaccttgggg aatttaatca aaaagggacg     360
atccgaacga aatacggaac aaaaacacaa tatattcaag ccattcaaac tgcccaagcc     420
gcagggatgc aagtatatgc ggatgttgta tttaatcata aggcaggggc tgacagtaca     480
gaatttgtcg atgcagttga ggtaaaccct tctaatcgaa atcaagaaac atctggcaca     540
tatcaaattc aagcatggac aaaatttgat tttcctggtc gtggaaacac atactccagc     600
ttcaaatggc gctggtacca ttttgatggt acggattggg acgaaagtcg taaattaaat     660
cgtatttaca aattccgcgg tacaggaaaa gcgtgggact gggaagtcga tacagaaaac     720
ggaaactatg attatttaat gttcgctgat ttagatatgg atcaccctga ggttgtgaca     780
gaattaaaaa actggggaac gtggtacgtc aatactacaa atatcgatgg attccgctta     840
gatgccgtaa aacatattaa atacagcttt ttccctgact ggctaacata tgtacgtaat     900
caaacaggaa aaaatttatt tgccgttggg gaattttgga gctatgacgt caataagctg     960
cataattaca ttacaaaaac aaatgggtcg atgtcattat ttgatgcacc cttgcataac    1020
aacttttata ccgcttccaa atcgagtgga tattttgaca tgcgttattt attgaataat    1080
acattaatga aagatcaacc ttcactcgct gtaacacttg tcgataacca cgacacgcaa    1140
ccagggcaat ctttacagtc atgggtcgaa ccttggttta aacagcttgc ttacgccttt    1200
attttaacaa gacaagaagg gtatccttgc gtattttacg gtgattatta tggaatccct    1260
aaatacaata tcccggggtt aaaaagtaaa atcgacccgc ttttaattgc tcgtcgtgat    1320
tacgcttatg gaacacaacg tgattacatt gatcatcaag acattatcgg atggacacga    1380
gaaggcattg atgcaaaacc gaactctgga ctggcggctt taattaccga cggtcctggt    1440
ggaagtaaat ggatgtatgt cggtaaaaag catgccggga aagtatttta tgatttaact    1500
ggaaatcgaa gtgacacagt aacgattaat gcggatggtt ggggagaatt taaagtaaac    1560
ggaggatccg tctcaatttg ggtggctaaa acgtcaaacg tcacatttac agtcaataac    1620
gccacaacaa caagcggaca aaacgtatat gttgtcggca acattccaga gctaggcaat    1680
tgtcgcacgg gttaa                                                     1695

<210> 120
<211> 564
<212> PRT
<213> Bacterial

<400> 120
Met Gln Thr Phe Ala Phe Leu Phe Tyr Ser Lys Lys Gly Trp Val Cys
 1               5                  10                  15
Met Asn Tyr Leu Lys Lys Val Trp Leu Tyr Tyr Ala Ile Val Ala Thr
            20                  25                  30
Leu Ile Ile Ser Phe Leu Thr Pro Phe Ser Thr Ala Gln Ala Asn Thr
        35                  40                  45
Ala Pro Val Asn Gly Thr Met Met Gln Tyr Phe Glu Trp Asp Leu Pro
    50                  55                  60
Asn Asp Gly Thr Leu Trp Thr Lys Val Lys Asn Glu Ala Thr Asn Leu
65                  70                  75                  80
Ser Ser Leu Gly Ile Thr Ala Leu Trp Leu Pro Pro Ala Tyr Lys Gly
                85                  90                  95
Thr Ser Gln Ser Asp Val Gly Tyr Gly Val Tyr Asp Leu Tyr Asp Leu
            100                 105                 110
Gly Glu Phe Asn Gln Lys Gly Thr Ile Arg Thr Lys Tyr Gly Thr Lys
        115                 120                 125
Thr Gln Tyr Ile Gln Ala Ile Gln Thr Ala Gln Ala Ala Gly Met Gln
    130                 135                 140
Val Tyr Ala Asp Val Val Phe Asn His Lys Ala Gly Ala Asp Ser Thr
145                 150                 155                 160
Glu Phe Val Asp Ala Val Glu Val Asn Pro Ser Asn Arg Asn Gln Glu
                165                 170                 175
Thr Ser Gly Thr Tyr Gln Ile Gln Ala Trp Thr Lys Phe Asp Phe Pro
            180                 185                 190
Gly Arg Gly Asn Thr Tyr Ser Ser Phe Lys Trp Arg Trp Tyr His Phe
        195                 200                 205
Asp Gly Thr Asp Trp Asp Glu Ser Arg Lys Leu Asn Arg Ile Tyr Lys
    210                 215                 220
Phe Arg Gly Thr Gly Lys Ala Trp Asp Trp Glu Val Asp Thr Glu Asn
225                 230                 235                 240
Gly Asn Tyr Asp Tyr Leu Met Phe Ala Asp Leu Asp Met Asp His Pro
                245                 250                 255
Glu Val Val Thr Glu Leu Lys Asn Trp Gly Thr Trp Tyr Val Asn Thr
            260                 265                 270
```

```
Thr Asn Ile Asp Gly Phe Arg Leu Asp Ala Val Lys His Ile Lys Tyr
        275                 280                 285
Ser Phe Phe Pro Asp Trp Leu Thr Tyr Val Arg Asn Gln Thr Gly Lys
    290                 295                 300
Asn Leu Phe Ala Val Gly Glu Phe Trp Ser Tyr Asp Val Asn Lys Leu
305                 310                 315                 320
His Asn Tyr Ile Thr Lys Thr Asn Gly Ser Met Ser Leu Phe Asp Ala
                325                 330                 335
Pro Leu His Asn Asn Phe Tyr Thr Ala Ser Lys Ser Ser Gly Tyr Phe
            340                 345                 350
Asp Met Arg Tyr Leu Leu Asn Asn Thr Leu Met Lys Asp Gln Pro Ser
        355                 360                 365
Leu Ala Val Thr Leu Val Asp Asn His Asp Thr Gln Pro Gly Gln Ser
    370                 375                 380
Leu Gln Ser Trp Val Glu Pro Trp Phe Lys Gln Leu Ala Tyr Ala Phe
385                 390                 395                 400
Ile Leu Thr Arg Gln Glu Gly Tyr Pro Cys Val Phe Tyr Gly Asp Tyr
                405                 410                 415
Tyr Gly Ile Pro Lys Tyr Asn Ile Pro Gly Leu Lys Ser Lys Ile Asp
            420                 425                 430
Pro Leu Leu Ile Ala Arg Arg Asp Tyr Ala Tyr Gly Thr Gln Arg Asp
        435                 440                 445
Tyr Ile Asp His Gln Asp Ile Ile Gly Trp Thr Arg Glu Gly Ile Asp
    450                 455                 460
Ala Lys Pro Asn Ser Gly Leu Ala Ala Leu Ile Thr Asp Gly Pro Gly
465                 470                 475                 480
Gly Ser Lys Trp Met Tyr Val Gly Lys Lys His Ala Gly Lys Val Phe
                485                 490                 495
Tyr Asp Leu Thr Gly Asn Arg Ser Asp Thr Val Thr Ile Asn Ala Asp
            500                 505                 510
Gly Trp Gly Glu Phe Lys Val Asn Gly Gly Ser Val Ser Ile Trp Val
        515                 520                 525
Ala Lys Thr Ser Asn Val Thr Phe Thr Val Asn Asn Ala Thr Thr Thr
    530                 535                 540
Ser Gly Gln Asn Val Tyr Val Val Gly Asn Ile Pro Glu Leu Gly Asn
545                 550                 555                 560
Cys Arg Thr Gly
```

<210> 121
<211> 1556
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 121

```
atgctcgccc tgtcgctcgg cggctgcggc atcgacgcgg gcccgacagg ccctcgcgtc      60
gtggagccgc tgccgcagcg ccccacgctt ccgcaggagt accgcgccag cggccacgcg     120
gccgccggcg acgtgttcgt gcacctgttc gagtggaagt ggccggacat cgcggaggaa     180
tgcgagaacg tgctggggcc ggcgggctac gaggcggtgc aggtgtcgcc gccgcaggag     240
cacctggtgc agcaggggcc gccgtggtgg cagcggtacc agccggtgag ctactcggtg     300
gcgctgagcc gcagcggcac gggcgtggag ttcagcaaca tgatcagccg gtgcaaggcc     360
gccggcgtgg acatctacgt ggacgccgtc atcaaccaca tgacggccgg tgcggggacg     420
gggagcaacg gcaccgccta caccaagtac aactaccccg gcctgtacgc gcaggcggac     480
tttcacccgc agtgcgcggt gggcgactac accagcgccg ccaacgtgca ggactgcgaa     540
ctgctggggc tggctgacct gaacaccggc gcggccggcg tgcagcagaa gatcgcggac     600
tacctggtct cgctggcgcg gctgggcgtg gcgggttttc gcatcgacgc cgccaagcac     660
atccagccgg tggaactgga cgccatcgtg gaccgcgtga accagacgct ggcggcggag     720
gggcgcccgc ttccctactg gttcgccgag gtgatcgaca acggcggcga gggggtgcgg     780
cgcgagcact actacggcct gggatacggc accggcggcg ccgcggacat cacggagttc     840
cgctacaagg gcgtgggcga caagttcctg ggcagcggcg gccagcggct ggtggacctg     900
aagaacttct cggcggtgac gtggaacctg atgccgtcgg acaaggccgt cgtctttctg     960
gagaaccacg atacgcagcg cggcggcggc atcggctacc gcgatggcac ggcgttccgg    1020
ctggccaacg tgtggatgct ggcgcagccg tacggctatc cgtcggtgat gtccagctac    1080
gcctttgacc gcacctcccc ctttggccgc gacgccggcc cgccctccga ggacggcgcg    1140
acgaaggacg tgacgtgcgc gcccacgctg gagacggcgg tgctgggcac ctgggtgtgc    1200
```

gagcaccgcg accccgtcat tcagcggatg gtgggctttc gccgcgcgat ggcgggcacg 1260
gacctgaacc gctggtggga caacggcggc aacgccattg ccttttcgcg cggggaccgg 1320
ggcttcgtcg ccatcagccg cgagccgaag gtgaccatgg cggccgtgcc cagcggactg 1380
tcccccggca cctactgcga cgtgctgacc ggcggcaagg tgggcaacgc ctgcgcggga 1440
accagcgtga cggtcgactc tcagggcgtg gtgcagctga gcatcgtcga gaactcggct 1500
ctggtgatcc acctcggggc caagctgtaa cggcgcgctg gcggatgtgc ggaggg 1556

<210> 122
<211> 517
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 122
Met Leu Ala Leu Ser Leu Gly Gly Cys Gly Ile Asp Ala Gly Pro Thr
 1   5   10   15
Gly Pro Arg Val Val Glu Pro Leu Pro Gln Arg Pro Thr Leu Pro Gln
            20   25   30
Glu Tyr Arg Ala Ser Gly His Ala Ala Ala Gly Asp Val Phe Val His
        35   40   45
Leu Phe Glu Trp Lys Trp Pro Asp Ile Ala Glu Glu Cys Glu Asn Val
    50   55   60
Leu Gly Pro Ala Gly Tyr Glu Ala Val Gln Val Ser Pro Pro Gln Glu
65   70   75   80
His Leu Val Gln Gln Gly Ala Pro Trp Trp Gln Arg Tyr Gln Pro Val
                85   90   95
Ser Tyr Ser Val Ala Leu Ser Arg Ser Gly Thr Gly Val Glu Phe Ser
            100   105   110
Asn Met Ile Ser Arg Cys Lys Ala Ala Gly Val Asp Ile Tyr Val Asp
        115   120   125
Ala Val Ile Asn His Met Thr Ala Gly Ala Gly Thr Gly Ser Asn Gly
    130   135   140
Thr Ala Tyr Thr Lys Tyr Asn Tyr Pro Gly Leu Tyr Ala Gln Ala Asp
145   150   155   160
Phe His Pro Gln Cys Ala Val Gly Asp Tyr Thr Ser Ala Ala Asn Val
                165   170   175
Gln Asp Cys Glu Leu Leu Gly Leu Ala Asp Leu Asn Thr Gly Ala Ala
            180   185   190
Gly Val Gln Gln Lys Ile Ala Asp Tyr Leu Val Ser Leu Ala Arg Leu
        195   200   205
Gly Val Ala Gly Phe Arg Ile Asp Ala Ala Lys His Ile Gln Pro Val
    210   215   220
Glu Leu Asp Ala Ile Val Asp Arg Val Asn Gln Thr Leu Ala Ala Glu
225   230   235   240
Gly Arg Pro Leu Pro Tyr Trp Phe Ala Glu Val Ile Asp Asn Gly Gly
                245   250   255
Glu Gly Val Arg Arg Glu His Tyr Tyr Gly Leu Gly Tyr Gly Thr Gly
            260   265   270
Gly Ala Ala Asp Ile Thr Glu Phe Arg Tyr Lys Gly Val Gly Asp Lys
        275   280   285
Phe Leu Gly Ser Gly Gly Gln Arg Leu Val Asp Leu Lys Asn Phe Ser
    290   295   300
Ala Val Thr Trp Asn Leu Met Pro Ser Asp Lys Ala Val Val Phe Leu
305   310   315   320
Glu Asn His Asp Thr Gln Arg Gly Gly Gly Ile Gly Tyr Arg Asp Gly
                325   330   335
Thr Ala Phe Arg Leu Ala Asn Val Trp Met Leu Ala Gln Pro Tyr Gly
            340   345   350
Tyr Pro Ser Val Met Ser Ser Tyr Ala Phe Asp Arg Thr Ser Pro Phe
        355   360   365
Gly Arg Asp Ala Gly Pro Pro Ser Glu Asp Gly Ala Thr Lys Asp Val
    370   375   380
Thr Cys Ala Pro Thr Leu Glu Thr Ala Val Leu Gly Thr Trp Val Cys
385   390   395   400
Glu His Arg Asp Pro Val Ile Gln Arg Met Val Gly Phe Arg Arg Ala
                405   410   415

```
Met Ala Gly Thr Asp Leu Asn Arg Trp Trp Asp Asn Gly Gly Asn Ala
            420                 425                 430
Ile Ala Phe Ser Arg Gly Asp Arg Gly Phe Val Ala Ile Ser Arg Glu
        435                 440                 445
Pro Lys Val Thr Met Ala Ala Val Pro Ser Gly Leu Ser Pro Gly Thr
    450                 455                 460
Tyr Cys Asp Val Leu Thr Gly Gly Lys Val Gly Asn Ala Cys Ala Gly
465                 470                 475                 480
Thr Ser Val Thr Val Asp Ser Gln Gly Val Val Gln Leu Ser Ile Val
                485                 490                 495
Glu Asn Ser Ala Leu Val Ile His Leu Gly Ala Lys Leu Arg Arg Ala
            500                 505                 510
Gly Gly Cys Ala Glu
        515
```

<210> 123
<211> 1770
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 123

```
atgccccagg ccattcgcac tttttcacgt tggacgttgt tcggcttaat cggcgttttt      60
ctgcttggtc tcgtcttttc tgtcccaccc cgggcaatcc aggcccagac aaccccggcc     120
cgtaccgtta tggttcacct cttcgagtgg aaatggaccg acatcgctaa agaatgcgag     180
aatttcctcg gaccgaaagg ctttgccgca atccaggtat cgccgcccca ggagcatgtc     240
caggggtcgc aatggtggac ccgctatcag ccggtcagct acaagatcga gagccgctcc     300
ggcacccggg ccgagttcgc caatatggtc tcgcgctgca aagccgtcgg ggtcgatatc     360
tatgtcgatg ccgtgatcaa ccatatgacg actgtcggct ccggcactgg tatggctgga     420
tcgacctaca ccagctacac ctatccgggg ctgtatcaga cccaggactt ccaccactgc     480
gggcgcaatg gcaacgatga tatcagcagc tacggcgatc gctgggaagt acaaaactgc     540
gaactgctca acctagccga cctcaacacc ggcgctgagt atgtccgggg taaactcgcc     600
gcctatatga acgatctgcg cggcctgggc gtcgccggat ttcggatcga tgccgccaag     660
cacatggata ccaacgacat caacaatatc gttggccgcc tgcccaacgc gccctacatc     720
taccaggaag tgatcgacca gggcggcgag ccaattaccg ccggcgaata cttccagaat     780
ggcgatgtga ccgagttcaa gtacagccgc gagatctcgc gcatgttcaa aaccggccag     840
ctgacccata tgagccagtt cggcactgcc tggggcttca tgtccagcga cctggcagta     900
gttttcaccg ataaccacga caccagcgc ggtcacggcg gcgccggcga tgtcttgacc     960
tacaaagatg gccagctgta caccctgggc aatatcttcg agctagcctg gccgtatggc    1020
tacccacagg tcatgtcgag ctacacgttc agcaacggcg accaggggcc gccatcgacc    1080
aatgtgtacg caaccacaac gcctgattgt ggcaacggcc gctgggtctg tgagcaccgc    1140
tggcgaggaa tcgccaacat ggtcgcgttc cgcaactaca ccgccccgac cttcagcacc    1200
agcaactggt ggagcaacgg caacaaccag atcgctttca gccgcgggac cctgggcttt    1260
gtggcgatca atcgggaagg tggcagcctg aaccgcacct tccaaaccgg cctgcccgtc    1320
ggcacctact gcgatgtcat tcacggcgat ttcaatgcca gcgccggcac ctgttccggc    1380
ccaactatcg ctgtcaacgg ctccggacag gcaaccatca cggtcaacgc gatggacgcg    1440
gtggcgatct acggcggagc caggctcgcc actccggcca gtgtcaacgt gacattcaac    1500
gaaaacgcca cgaccacctg ggggcagaat gtgtatatcg tcggcaacgt cgccgccctg    1560
ggcagctgga acgcaggcag cgcggtctta ctctcctccg ctaactaccc aatctggagc    1620
aagaccatcg ccctgccagc caacaccgcc attgagtaca agtacatcaa aaaggatggc    1680
gcgggcaatg tggtgtggga aagcggcgcc aaccgcgtct ttaccacccc cggcagcggc    1740
agtgccacgc gcaacgatac ctggaaatag                                     1770
```

<210> 124
<211> 589
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 124

```
Met Pro Gln Ala Ile Arg Thr Phe Ser Arg Trp Thr Leu Phe Gly Leu
 1               5                  10                  15
Ile Gly Val Phe Leu Leu Gly Leu Val Phe Ser Val Pro Pro Arg Ala
            20                  25                  30
```

```
Ile Gln Ala Gln Thr Thr Pro Ala Arg Thr Val Met Val His Leu Phe
        35                  40                  45
Glu Trp Lys Trp Thr Asp Ile Ala Lys Glu Cys Glu Asn Phe Leu Gly
    50                  55                  60
Pro Lys Gly Phe Ala Ala Ile Gln Val Ser Pro Pro Gln Glu His Val
65                  70                  75                  80
Gln Gly Ser Gln Trp Trp Thr Arg Tyr Gln Pro Val Ser Tyr Lys Ile
                85                  90                  95
Glu Ser Arg Ser Gly Thr Arg Ala Glu Phe Ala Asn Met Val Ser Arg
            100                 105                 110
Cys Lys Ala Val Gly Val Asp Ile Tyr Val Asp Ala Val Ile Asn His
        115                 120                 125
Met Thr Thr Val Gly Ser Gly Thr Gly Met Ala Gly Ser Thr Tyr Thr
    130                 135                 140
Ser Tyr Thr Tyr Pro Gly Leu Tyr Gln Thr Gln Asp Phe His His Cys
145                 150                 155                 160
Gly Arg Asn Gly Asn Asp Asp Ile Ser Ser Tyr Gly Asp Arg Trp Glu
                165                 170                 175
Val Gln Asn Cys Glu Leu Leu Asn Leu Ala Asp Leu Asn Thr Gly Ala
            180                 185                 190
Glu Tyr Val Arg Gly Lys Leu Ala Ala Tyr Met Asn Asp Leu Arg Gly
        195                 200                 205
Leu Gly Val Ala Gly Phe Arg Ile Asp Ala Ala Lys His Met Asp Thr
    210                 215                 220
Asn Asp Ile Asn Asn Ile Val Gly Arg Leu Pro Asn Ala Pro Tyr Ile
225                 230                 235                 240
Tyr Gln Glu Val Ile Asp Gln Gly Gly Glu Pro Ile Thr Ala Gly Glu
                245                 250                 255
Tyr Phe Gln Asn Gly Asp Val Thr Glu Phe Lys Tyr Ser Arg Glu Ile
            260                 265                 270
Ser Arg Met Phe Lys Thr Gly Gln Leu Thr His Met Ser Gln Phe Gly
        275                 280                 285
Thr Ala Trp Gly Phe Met Ser Ser Asp Leu Ala Val Val Phe Thr Asp
    290                 295                 300
Asn His Asp Asn Gln Arg Gly His Gly Gly Ala Gly Asp Val Leu Thr
305                 310                 315                 320
Tyr Lys Asp Gly Gln Leu Tyr Thr Leu Gly Asn Ile Phe Glu Leu Ala
                325                 330                 335
Trp Pro Tyr Gly Tyr Pro Gln Val Met Ser Ser Tyr Thr Phe Ser Asn
            340                 345                 350
Gly Asp Gln Gly Pro Pro Ser Thr Asn Val Tyr Ala Thr Thr Thr Pro
        355                 360                 365
Asp Cys Gly Asn Gly Arg Trp Val Cys Glu His Arg Trp Arg Gly Ile
    370                 375                 380
Ala Asn Met Val Ala Phe Arg Asn Tyr Thr Ala Pro Thr Phe Ser Thr
385                 390                 395                 400
Ser Asn Trp Trp Ser Asn Gly Asn Asn Gln Ile Ala Phe Ser Arg Gly
                405                 410                 415
Thr Leu Gly Phe Val Ala Ile Asn Arg Glu Gly Gly Ser Leu Asn Arg
            420                 425                 430
Thr Phe Gln Thr Gly Leu Pro Val Gly Thr Tyr Cys Asp Val Ile His
        435                 440                 445
Gly Asp Phe Asn Ala Ser Ala Gly Thr Cys Ser Gly Pro Thr Ile Ala
    450                 455                 460
Val Asn Gly Ser Gly Gln Ala Thr Ile Thr Val Asn Ala Met Asp Ala
465                 470                 475                 480
Val Ala Ile Tyr Gly Gly Ala Arg Leu Ala Thr Pro Ala Ser Val Asn
                485                 490                 495
Val Thr Phe Asn Glu Asn Ala Thr Thr Thr Trp Gly Gln Asn Val Tyr
            500                 505                 510
Ile Val Gly Asn Val Ala Ala Leu Gly Ser Trp Asn Ala Gly Ser Ala
        515                 520                 525
Val Leu Leu Ser Ser Ala Asn Tyr Pro Ile Trp Ser Lys Thr Ile Ala
    530                 535                 540
Leu Pro Ala Asn Thr Ala Ile Glu Tyr Lys Tyr Ile Lys Lys Asp Gly
545                 550                 555                 560
Ala Gly Asn Val Val Trp Glu Ser Gly Ala Asn Arg Val Phe Thr Thr
                565                 570                 575
```

Pro Gly Ser Gly Ser Ala Thr Arg Asn Asp Thr Trp Lys
580 585

<210> 125
<211> 1395
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 125
gtggtgcaca tgaagttgaa gtaccttgcc ttagttttgt tggctgtggc ttcgataggc 60
ctactctcga ctccagtggg tgctgccaag tactccgaac tcgaagaggg cggtgttata 120
atgcaggcct tctactggga tgttcccgga gggggaatct ggtgggacac cataagacag 180
aaaatcccgg agtggtacga cgctggaatc tcggcgatat ggattcctcc agctagcaaa 240
gggatgggcg gtggttattc catgggctac gatccctacg atttctttga cctcggcgag 300
tactatcaga agggaacagt tgagacgcgc ttcggctcaa aggaggaact ggtgaacatg 360
ataaacaccg cacactccta tggcataaag gtgatagcgg acatagtcat aaaccaccgc 420
gccggtggag accttgagtg gaacccttt gtaaacaact atacttggac agacttctcc 480
aaggtcgcct ccggtaaata cacggccaac taccttgact tccacccaaa cgaggtcaag 540
tgctgcgatg agggtacatt tggtgacttt ccggacatcg cccacgagaa gagctgggat 600
cagtactggc tctgggcaag caatgagagc tacgccgcat atctccggag catagggatc 660
gatgcatggc gtttcgacta cgtcaaaggt tacggagcgt gggttgttaa tgactggctc 720
agctggtggg gaggctgggc cgttggagag tactgggaca cgaacgttga tgcactcctt 780
aactgggcat acgacagcgg tgccaaggtc tttgacttcc cgctctacta caagatggac 840
gaagcctttg acaacaccaa catccccgct ttggtttacg ccctccagaa cggaggaaca 900
gtcgtttccc gcgatccctt caaggcagta actttcgttg ccaaccacga tacagatata 960
atctggaaca agtatccggc ttatgcgttc atccttacct atgagggaca gcctgttata 1020
ttttaccgcg actacgagga gtggctcaac aaggataagc ttaacaacct tatctggata 1080
cacgagcacc ttgccggagg aagtaccaag atcctctact acgataacga tgagctaata 1140
ttcatgaggg agggctacgg gagcaagccg ggcctcataa cctacataaa cctcggaaac 1200
gactgggccg agcgctgggt gaacgtcggc tcaaagtttg ccggctacac aatccatgaa 1260
tacacaggca atctcggtgg ctgggttgac aggtgggttc agtacgatgg atgggttaaa 1320
ctgacggcac ctcctcatga tccagccaac ggatattacg gctactcagt ctggagctac 1380
gcaggcgtcg gatga 1395

<210> 126
<211> 464
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 126
Val Val His Met Lys Leu Lys Tyr Leu Ala Leu Val Leu Leu Ala Val
1 5 10 15
Ala Ser Ile Gly Leu Leu Ser Thr Pro Val Gly Ala Ala Lys Tyr Ser
20 25 30
Glu Leu Glu Glu Gly Gly Val Ile Met Gln Ala Phe Tyr Trp Asp Val
35 40 45
Pro Gly Gly Gly Ile Trp Trp Asp Thr Ile Arg Gln Lys Ile Pro Glu
50 55 60
Trp Tyr Asp Ala Gly Ile Ser Ala Ile Trp Ile Pro Pro Ala Ser Lys
65 70 75 80
Gly Met Gly Gly Gly Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Phe Phe
85 90 95
Asp Leu Gly Glu Tyr Tyr Gln Lys Gly Thr Val Glu Thr Arg Phe Gly
100 105 110
Ser Lys Glu Glu Leu Val Asn Met Ile Asn Thr Ala His Ser Tyr Gly
115 120 125
Ile Lys Val Ile Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly Asp
130 135 140
Leu Glu Trp Asn Pro Phe Val Asn Asn Tyr Thr Trp Thr Asp Phe Ser
145 150 155 160
Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr Leu Asp Phe His Pro
165 170 175

```
Asn Glu Val Lys Cys Cys Asp Glu Gly Thr Phe Gly Asp Phe Pro Asp
            180                 185                 190
Ile Ala His Glu Lys Ser Trp Asp Gln Tyr Trp Leu Trp Ala Ser Asn
        195                 200                 205
Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly Ile Asp Ala Trp Arg
    210                 215                 220
Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp Val Val Asn Asp Trp Leu
225                 230                 235                 240
Ser Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn Val
                245                 250                 255
Asp Ala Leu Leu Asn Trp Ala Tyr Asp Ser Gly Ala Lys Val Phe Asp
            260                 265                 270
Phe Pro Leu Tyr Tyr Lys Met Asp Glu Ala Phe Asp Asn Thr Asn Ile
        275                 280                 285
Pro Ala Leu Val Tyr Ala Leu Gln Asn Gly Gly Thr Val Val Ser Arg
    290                 295                 300
Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Asp Ile
305                 310                 315                 320
Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile Leu Thr Tyr Glu Gly
                325                 330                 335
Gln Pro Val Ile Phe Tyr Arg Asp Tyr Glu Glu Trp Leu Asn Lys Asp
            340                 345                 350
Lys Leu Asn Asn Leu Ile Trp Ile His Glu His Leu Ala Gly Gly Ser
        355                 360                 365
Thr Lys Ile Leu Tyr Tyr Asp Asn Asp Glu Leu Ile Phe Met Arg Glu
    370                 375                 380
Gly Tyr Gly Ser Lys Pro Gly Leu Ile Thr Tyr Ile Asn Leu Gly Asn
385                 390                 395                 400
Asp Trp Ala Glu Arg Trp Val Asn Val Gly Ser Lys Phe Ala Gly Tyr
                405                 410                 415
Thr Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp Val Asp Arg Trp
            420                 425                 430
Val Gln Tyr Asp Gly Trp Val Lys Leu Thr Ala Pro Pro His Asp Pro
        435                 440                 445
Ala Asn Gly Tyr Tyr Gly Tyr Ser Val Trp Ser Tyr Ala Gly Val Gly
    450                 455                 460
```

<210> 127
<211> 1848
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 127

```
gtgtgcatga attatttgaa aaaagtgtgg ttgtattacg ctatcgtcgc taccttaatc      60
atttactttc ttacgccctt ttcaactgca caagccaaca ctgcaccagt caacggaacg     120
atgatgcaat atttcgaatg ggatttaccg aatgatggca cactttggac gaaagtaaaa     180
aacgaagcaa gcagtctttc ttctttaggt attactgcgt tatggttacc acctgcatac     240
aaaggaacga gccaagggga tgtcgggtat ggcgtgtacg atttgtatga cttaggagaa     300
tttaatcaaa aagggacgat tcgaacgaaa tacggaacaa aaacgcaata tttacaagcc     360
attcaagcgg caaaaagcgc tggcatgcaa gtatacgctg atgtcgtatt taatcacaag     420
gcgggggcag atagtacaga atgggttgac gcagtcgaag tgaatccttc taatcgaaac     480
caagaaacat ctggcacata tcaaattcaa gcatggacaa aatttgattt ccctggccgt     540
gggaacacat actcaagctt taaatggcga tggtatcatt ttgacggtac ggattgggat     600
gaaagccgaa aactaaatcg tatttacaaa tttcgtggca caggaaaagc atgggattgg     660
gaagtagaca cagagaacgg aaactatgac tacttaatgt ttgctgattt agatatggat     720
caccctgaag tcgtgacaga gctaaaaaac tggggaacat ggtacgtcaa tacgacaaat     780
gtcgatgggt ttcgcttaga tgcagtaaag catattaaat atagcttctt cccagattgg     840
ttaacacatg tgcgttcaca aacacgaaaa aatctttttg cagtaggaga attttggagc     900
tacgatgtca ataaactgca taactacatt acaaaaacaa gtggaaccat gtcgttattt     960
gatgcgccac ttcataacaa cttttacact gcttcaaaat ctagcgggta ttttgacatg    1020
cgctatttgt taaataatac gttgatgaaa gaccagcctt ctcttgcggt cacactcgtt    1080
gataatcatg acacgcaacc gggacaatct ttacaatcat gggtagagcc ttggtttaag    1140
ccgcttgctt atgcctttat tttgacaaga caagaaggat atccttgcgt attttacggc    1200
gactattacg gcatccctaa atacaacatt ccgggattga aaagtaaaat cgatccgctt    1260
ctcattgccc gtagagacta cgcatacgga acacaacgtg attatattga ccatcaagac    1320
```

attattggat ggacacggga aggaattgac tcaaaaccga actctggact tgcggcttta 1380
attactgacg gccctggtgg aagtaaatgg atgtatgtag gtaaaaagca tgctggaaaa 1440
gtgttttacg atctcactgg aaatcgaagc gatacggtaa cgattaatgc agacggctgg 1500
ggagagttta aagtaaacgg tggctccgtt tccatttggg ttgccaaaac atcacaagtc 1560
acgtttaccg tcaacaatgc gacaacgata agcggacaaa atgtgtatgt cgttggtaac 1620
attccagagc tcggaaattg gaacacagca aacgcaatca aaatgacccc atcttcttat 1680
ccaacgtgga aagcaaccat tgctcttcca caaggaaaag ccattgaatt taaatttatt 1740
aaaaaagacc aatcgggaaa tgttgtttgg gaaagcattc caaaccgaac atacaccgtt 1800
ccatttttat caacaggctc atatacagct agttggaatg taccttaa 1848

<210> 128
<211> 615
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 128
Val Cys Met Asn Tyr Leu Lys Lys Val Trp Leu Tyr Tyr Ala Ile Val
 1   5   10   15
Ala Thr Leu Ile Ile Tyr Phe Leu Thr Pro Phe Ser Thr Ala Gln Ala
            20   25   30
Asn Thr Ala Pro Val Asn Gly Thr Met Met Gln Tyr Phe Glu Trp Asp
        35   40   45
Leu Pro Asn Asp Gly Thr Leu Trp Thr Lys Val Lys Asn Glu Ala Ser
    50   55   60
Ser Leu Ser Ser Leu Gly Ile Thr Ala Leu Trp Leu Pro Pro Ala Tyr
65   70   75   80
Lys Gly Thr Ser Gln Gly Asp Val Gly Tyr Gly Val Tyr Asp Leu Tyr
                85   90   95
Asp Leu Gly Glu Phe Asn Gln Lys Gly Thr Ile Arg Thr Lys Tyr Gly
            100   105   110
Thr Lys Thr Gln Tyr Leu Gln Ala Ile Gln Ala Ala Lys Ser Ala Gly
        115   120   125
Met Gln Val Tyr Ala Asp Val Val Phe Asn His Lys Ala Gly Ala Asp
    130   135   140
Ser Thr Glu Trp Val Asp Ala Val Glu Val Asn Pro Ser Asn Arg Asn
145   150   155   160
Gln Glu Thr Ser Gly Thr Tyr Gln Ile Gln Ala Trp Thr Lys Phe Asp
                165   170   175
Phe Pro Gly Arg Gly Asn Thr Tyr Ser Ser Phe Lys Trp Arg Trp Tyr
            180   185   190
His Phe Asp Gly Thr Asp Trp Asp Glu Ser Arg Lys Leu Asn Arg Ile
        195   200   205
Tyr Lys Phe Arg Gly Thr Gly Lys Ala Trp Asp Trp Glu Val Asp Thr
    210   215   220
Glu Asn Gly Asn Tyr Asp Tyr Leu Met Phe Ala Asp Leu Asp Met Asp
225   230   235   240
His Pro Glu Val Val Thr Glu Leu Lys Asn Trp Gly Thr Trp Tyr Val
                245   250   255
Asn Thr Thr Asn Val Asp Gly Phe Arg Leu Asp Ala Val Lys His Ile
            260   265   270
Lys Tyr Ser Phe Phe Pro Asp Trp Leu Thr His Val Arg Ser Gln Thr
        275   280   285
Arg Lys Asn Leu Phe Ala Val Gly Glu Phe Trp Ser Tyr Asp Val Asn
    290   295   300
Lys Leu His Asn Tyr Ile Thr Lys Thr Ser Gly Thr Met Ser Leu Phe
305   310   315   320
Asp Ala Pro Leu His Asn Asn Phe Tyr Thr Ala Ser Lys Ser Ser Gly
                325   330   335
Tyr Phe Asp Met Arg Tyr Leu Leu Asn Asn Thr Leu Met Lys Asp Gln
            340   345   350
Pro Ser Leu Ala Val Thr Leu Val Asp Asn His Asp Thr Gln Pro Gly
        355   360   365
Gln Ser Leu Gln Ser Trp Val Glu Pro Trp Phe Lys Pro Leu Ala Tyr
    370   375   380
Ala Phe Ile Leu Thr Arg Gln Glu Gly Tyr Pro Cys Val Phe Tyr Gly

385 390 395 400
Asp Tyr Tyr Gly Ile Pro Lys Tyr Asn Ile Pro Gly Leu Lys Ser Lys
405 410 415
Ile Asp Pro Leu Leu Ile Ala Arg Arg Asp Tyr Ala Tyr Gly Thr Gln
420 425 430
Arg Asp Tyr Ile Asp His Gln Asp Ile Ile Gly Trp Thr Arg Glu Gly
435 440 445
Ile Asp Ser Lys Pro Asn Ser Gly Leu Ala Ala Leu Ile Thr Asp Gly
450 455 460
Pro Gly Gly Ser Lys Trp Met Tyr Val Gly Lys Lys His Ala Gly Lys
465 470 475 480
Val Phe Tyr Asp Leu Thr Gly Asn Arg Ser Asp Thr Val Thr Ile Asn
485 490 495
Ala Asp Gly Trp Gly Glu Phe Lys Val Asn Gly Gly Ser Val Ser Ile
500 505 510
Trp Val Ala Lys Thr Ser Gln Val Thr Phe Thr Val Asn Asn Ala Thr
515 520 525
Thr Ile Ser Gly Gln Asn Val Tyr Val Val Gly Asn Ile Pro Glu Leu
530 535 540
Gly Asn Trp Asn Thr Ala Asn Ala Ile Lys Met Thr Pro Ser Ser Tyr
545 550 555 560
Pro Thr Trp Lys Ala Thr Ile Ala Leu Pro Gln Gly Lys Ala Ile Glu
565 570 575
Phe Lys Phe Ile Lys Lys Asp Gln Ser Gly Asn Val Val Trp Glu Ser
580 585 590
Ile Pro Asn Arg Thr Tyr Thr Val Pro Phe Leu Ser Thr Gly Ser Tyr
595 600 605
Thr Ala Ser Trp Asn Val Pro
610 615

<210> 129
<211> 1854
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 129
atgcgttgcc gccgtggcag ggacgggtgt tggtgcgggc ggcgtaatgc gctgccgcga 60
cacccgcgtg aacaaaataa tatgaattat ttgaatagga tgggggtgtc aagaatgaca 120
aaatctcgag agttgcggtg ttcatggaaa gtatttgttg ttgggtgcct gttgtggatg 180
gcttggggat cttccgcgtc cgccggcgta ttgatgcaag gcttctactg ggacgccagt 240
accgggacca gtgattcgtg gtggacgcat ttggccaagc aagccaacgg tctaaaacgg 300
gcggggttca ccgccgtatg gattcctccg gtgcttaaag gggcttcagg gggctattcc 360
aacgggtacg atccctttga cgactatgat atcggaagca aggaccagaa aggtaccgtg 420
gcgacgcgat gggggacgcg agaagaactg caacgtgccg tggccgtgat gcgcgcgaac 480
ggtctggatg tgtatgtgga tctggtgctg aaccaccgca acggggacga cgggaattgg 540
aattttcatt acaaagatgc gtacggcaaa gtgggttacg ggcggtttca aaaggggttt 600
tacgattttc accccaacta caacattcag gatgccaatg ttcccaacga ggattccagc 660
ttcgggcgcg atttagccca tgacaatccg tatgtggccg atggactgaa ggctgcaggc 720
gattggctga ccaaagccct cgatgttcag ggatatcgtc tggattacgt gaaaggcatc 780
agctacacct tcctgaaaag ttatctgtcc tatggggcca tgaacggaaa atttgccgtc 840
ggtgagtact gggatgccaa ccgggatacg ttgaactggt gggcgaacac ggcgatggaa 900
gggcgggccc atgtgtttga ttttgcgttg cgcgaggagc tgaaaaacat gtgcaatgcg 960
gacgggtact acgacatgcg tcgattggac cacgcgggtc tggtcggaat cgacccgtgg 1020
aaggcggtga cgtttgtcga aaaccatgat acggatcggc acgaccccat ctacaataac 1080
aagcatttgg cgtatgccta catcttgacg tcggaagggt atccgacggt gttctggaag 1140
gattactacc aatacggaat gaagccgatc atcgacaacc tcatttggat ccacgaacac 1200
attgcgtacg gaacgaccca agagcgttgg aaagacgaag atgtctttgt gtatgagcgg 1260
accggaggca agcggctatt ggtggggctt aacgacaatc gcgccaccag caaaacggtc 1320
accgtacaga ccggctttgg tgccaacgtg gccttgcacg actacaccgg caacggcccc 1380
gatctccgta ccgacgccta cggtcgggta accttgacca ttcctgcaaa cgggtacgtg 1440
gcctattccg ttccgggcat ctccggatcc tttgtgccgg tcgagaaaac cgtgacgcag 1500
gagtttgccg gggcgtccga cttggatatt cgtccggccg ataacacgca atttgtgcag 1560
gtcgggcgga tatacgccaa ggcaaacaag ccggttacag cggaattgta ttgggatgcc 1620
aaagactgga cgacctccac gtcgattctc ctagaagtgc gttcggcttc gggaacgctc 1680
atcacgacaa agaccgtgac ccaattgtcg tcccagggta cccgcgtttc cttcacgcct 1740

tcggctaccg gatggtacgt cttttccatt cgaagctata acacgccttc gacgaaccca 1800
aagccggcct actggttaaa ggtaacgtat acggcgccgc aattgcttca gtaa 1854

<210> 130
<211> 617
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 130

```
Met Arg Cys Arg Arg Gly Arg Asp Gly Cys Trp Cys Gly Arg Arg Asn
 1               5                  10                  15
Ala Leu Pro Arg His Pro Arg Glu Gln Asn Asn Met Asn Tyr Leu Asn
            20                  25                  30
Arg Met Gly Val Ser Arg Met Thr Lys Ser Arg Glu Leu Arg Cys Ser
        35                  40                  45
Trp Lys Val Phe Val Val Gly Cys Leu Leu Trp Met Ala Trp Gly Ser
    50                  55                  60
Ser Ala Ser Ala Gly Val Leu Met Gln Gly Phe Tyr Trp Asp Ala Ser
65                  70                  75                  80
Thr Gly Thr Ser Asp Ser Trp Trp Thr His Leu Ala Lys Gln Ala Asn
                85                  90                  95
Gly Leu Lys Arg Ala Gly Phe Thr Ala Val Trp Ile Pro Pro Val Leu
            100                 105                 110
Lys Gly Ala Ser Gly Gly Tyr Ser Asn Gly Tyr Asp Pro Phe Asp Asp
        115                 120                 125
Tyr Asp Ile Gly Ser Lys Asp Gln Lys Gly Thr Val Ala Thr Arg Trp
    130                 135                 140
Gly Thr Arg Glu Glu Leu Gln Arg Ala Val Ala Val Met Arg Ala Asn
145                 150                 155                 160
Gly Leu Asp Val Tyr Val Asp Leu Val Leu Asn His Arg Asn Gly Asp
                165                 170                 175
Asp Gly Asn Trp Asn Phe His Tyr Lys Asp Ala Tyr Gly Lys Val Gly
            180                 185                 190
Tyr Gly Arg Phe Gln Lys Gly Phe Tyr Asp Phe His Pro Asn Tyr Asn
        195                 200                 205
Ile Gln Asp Ala Asn Val Pro Asn Glu Asp Ser Ser Phe Gly Arg Asp
    210                 215                 220
Leu Ala His Asp Asn Pro Tyr Val Ala Asp Gly Leu Lys Ala Ala Gly
225                 230                 235                 240
Asp Trp Leu Thr Lys Ala Leu Asp Val Gln Gly Tyr Arg Leu Asp Tyr
                245                 250                 255
Val Lys Gly Ile Ser Tyr Thr Phe Leu Lys Ser Tyr Leu Ser Tyr Gly
            260                 265                 270
Ala Met Asn Gly Lys Phe Ala Val Gly Glu Tyr Trp Asp Ala Asn Arg
        275                 280                 285
Asp Thr Leu Asn Trp Trp Ala Asn Thr Ala Met Glu Gly Arg Ala His
    290                 295                 300
Val Phe Asp Phe Ala Leu Arg Glu Glu Leu Lys Asn Met Cys Asn Ala
305                 310                 315                 320
Asp Gly Tyr Tyr Asp Met Arg Arg Leu Asp His Ala Gly Leu Val Gly
                325                 330                 335
Ile Asp Pro Trp Lys Ala Val Thr Phe Val Glu Asn His Asp Thr Asp
            340                 345                 350
Arg His Asp Pro Ile Tyr Asn Asn Lys His Leu Ala Tyr Ala Tyr Ile
        355                 360                 365
Leu Thr Ser Glu Gly Tyr Pro Thr Val Phe Trp Lys Asp Tyr Tyr Gln
    370                 375                 380
Tyr Gly Met Lys Pro Ile Ile Asp Asn Leu Ile Trp Ile His Glu His
385                 390                 395                 400
Ile Ala Tyr Gly Thr Thr Gln Glu Arg Trp Lys Asp Glu Asp Val Phe
                405                 410                 415
Val Tyr Glu Arg Thr Gly Gly Lys Arg Leu Leu Val Gly Leu Asn Asp
            420                 425                 430
Asn Arg Ala Thr Ser Lys Thr Val Thr Val Gln Thr Gly Phe Gly Ala
        435                 440                 445
```

Asn Val Ala Leu His Asp Tyr Thr Gly Asn Gly Pro Asp Leu Arg Thr
450 455 460
Asp Ala Tyr Gly Arg Val Thr Leu Thr Ile Pro Ala Asn Gly Tyr Val
465 470 475 480
Ala Tyr Ser Val Pro Gly Ile Ser Gly Ser Phe Val Pro Val Glu Lys
485 490 495
Thr Val Thr Gln Glu Phe Ala Gly Ala Ser Asp Leu Asp Ile Arg Pro
500 505 510
Ala Asp Asn Thr Gln Phe Val Gln Val Gly Arg Ile Tyr Ala Lys Ala
515 520 525
Asn Lys Pro Val Thr Ala Glu Leu Tyr Trp Asp Ala Lys Asp Trp Thr
530 535 540
Thr Ser Thr Ser Ile Leu Leu Glu Val Arg Ser Ala Ser Gly Thr Leu
545 550 555 560
Ile Thr Thr Lys Thr Val Thr Gln Leu Ser Ser Gln Gly Thr Arg Val
565 570 575
Ser Phe Thr Pro Ser Ala Thr Gly Trp Tyr Val Phe Ser Ile Arg Ser
580 585 590
Tyr Asn Thr Pro Ser Thr Asn Pro Lys Pro Ala Tyr Trp Leu Lys Val
595 600 605
Thr Tyr Thr Ala Pro Gln Leu Leu Gln
610 615

<210> 131
<211> 1881
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 131
atgccgcagc tttacccatt gccgccgcgc tggcggcgcg cggcccggca gggcctggcc 60
gccttgacgc tggccaccac ggccctgggc atctcgacgg cccaggccca gagtgcaccg 120
cgcacggcct tcgtgcatct gttcgaatgg aagtggaccg acatcgcgcg cgagtgcgag 180
accttcctcg ggcccaaggg cttcgcggcg gtgcaggtgt cgcccccgaa cgagcacaac 240
tgggtgacca gcggtgatgg tgcaccttat ccgtggtgga tgcgctacca gccggtgagc 300
tacagcctgg accgcagccg cagcggcacg cgcgccgagt tccaggacat ggtcaaccga 360
tgcaatgccg tgggcgtggg catctacgtg gacgccgtga tcaatcacat gtccggcggc 420
acgggcggca cctcgagcgc tgggcgcagc tggagctatc acaactaccc tgggctctat 480
ggccccaacg acttccacca gccggtgtgc agcatcacca actacgggga tgcgaacaat 540
gtgcagcgtt gcgagctctc gggcttgcag gacctggaca ctgggagcgc ttatgtgcgc 600
ggcaagatcg ccgactatct ggtggatctg gtcaacatgg gggtcaaggg cttccgggtg 660
gatgcggcca agcacatcag cccgaccgac ctgggcgcca tcatcgatgc ggtcaacagc 720
cgcaccggcg cgaaccgccc tttctggttt ctggaggtga ttggcgcggc cggcgaggca 780
gtgcagccga accagtactt ctcgctcggc ggcggccagg tcaccgtgac cgagttcaac 840
tatgggaagc aaatcttcgg caagttcgcc ggtggcggcc gtctggccga gctgcgcagc 900
ttcggtgaaa cctggggcct gatgcccagc agcaaagcga ttgctttcat cgacaaccac 960
gacaagcagc gcggtcatgg cggcggtggc aactatctga cctaccacca tggctcgacc 1020
tacgatctgg ccaacatctt catgctggct tggccttatg gctacccggc gctgatgtcc 1080
agctatgcct tcaaccgcag cacggcctac gacacgagct ttggcccgcc acacgacagt 1140
ggtggcgcca cccgtggccc ctgggatggt ggcggcagcc agccggcctg cttcaaccag 1200
agcatcggtg gctgggtgtg tgagcaccgc tggcggggca tcgccaatat ggtggccttc 1260
cgcaacgcca cgctgcccaa ctggaccgtg accgactggt gggacaacgg caacaaccag 1320
atcgctttcg ggcggggtga caagggcttc gtggtgatca accgcgaaga cgccgcgctg 1380
acgcgcaact tcaagaccag cctgccagcc ggccagtact gcgatgtcat ctccggggac 1440
ttcaacaatg gtcagtgcac gggccatgtg gtgacggtcg atgccggcgg ctacgtgacg 1500
ctgacggccg ggcccaatgg tgcggcggcc atccacgtgg gcgcccgtct ggacggcgcc 1560
tctcagccgc cgacgaccgc ctcggtgacg ttcaacgcgt cggccgatac cttttgggga 1620
cagaacctgt tcgtcgtggg caaccacagc gcactgggca actggtcgcc ggcggccgcc 1680
aggccgatga cttggatttc cggttcgggc acgcgcggga actggcgcgc ggtgctcaat 1740
ttgccggcca ataccaccta ccaatacaag ttcatcaaga aggacggggc tggaaacgtg 1800
gtttgggagg gcggtggcaa tcgcgtcgtg accacgccgt ctgggggcgg atcggtgagc 1860
acgggcggca attggcagta g 1881

<210> 132
<211> 626
<212> PRT

<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 132

```
Met Pro Gln Leu Tyr Pro Leu Pro Pro Arg Trp Arg Arg Ala Ala Arg
 1               5                  10                  15
Gln Gly Leu Ala Ala Leu Thr Leu Ala Thr Thr Ala Leu Gly Ile Ser
            20                  25                  30
Thr Ala Gln Ala Gln Ser Ala Pro Arg Thr Ala Phe Val His Leu Phe
        35                  40                  45
Glu Trp Lys Trp Thr Asp Ile Ala Arg Glu Cys Glu Thr Phe Leu Gly
    50                  55                  60
Pro Lys Gly Phe Ala Ala Val Gln Val Ser Pro Pro Asn Glu His Asn
65                  70                  75                  80
Trp Val Thr Ser Gly Asp Gly Ala Pro Tyr Pro Trp Trp Met Arg Tyr
                85                  90                  95
Gln Pro Val Ser Tyr Ser Leu Asp Arg Ser Arg Ser Gly Thr Arg Ala
            100                 105                 110
Glu Phe Gln Asp Met Val Asn Arg Cys Asn Ala Val Gly Val Gly Ile
        115                 120                 125
Tyr Val Asp Ala Val Ile Asn His Met Ser Gly Gly Thr Gly Gly Thr
    130                 135                 140
Ser Ser Ala Gly Arg Ser Trp Ser Tyr His Asn Tyr Pro Gly Leu Tyr
145                 150                 155                 160
Gly Pro Asn Asp Phe His Gln Pro Val Cys Ser Ile Thr Asn Tyr Gly
                165                 170                 175
Asp Ala Asn Asn Val Gln Arg Cys Glu Leu Ser Gly Leu Gln Asp Leu
            180                 185                 190
Asp Thr Gly Ser Ala Tyr Val Arg Gly Lys Ile Ala Asp Tyr Leu Val
        195                 200                 205
Asp Leu Val Asn Met Gly Val Lys Gly Phe Arg Val Asp Ala Ala Lys
    210                 215                 220
His Ile Ser Pro Thr Asp Leu Gly Ala Ile Ile Asp Ala Val Asn Ser
225                 230                 235                 240
Arg Thr Gly Ala Asn Arg Pro Phe Trp Phe Leu Glu Val Ile Gly Ala
                245                 250                 255
Ala Gly Glu Ala Val Gln Pro Asn Gln Tyr Phe Ser Leu Gly Gly Gly
            260                 265                 270
Gln Val Thr Val Thr Glu Phe Asn Tyr Gly Lys Gln Ile Phe Gly Lys
        275                 280                 285
Phe Ala Gly Gly Gly Arg Leu Ala Glu Leu Arg Ser Phe Gly Glu Thr
    290                 295                 300
Trp Gly Leu Met Pro Ser Ser Lys Ala Ile Ala Phe Ile Asp Asn His
305                 310                 315                 320
Asp Lys Gln Arg Gly His Gly Gly Gly Gly Asn Tyr Leu Thr Tyr His
                325                 330                 335
His Gly Ser Thr Tyr Asp Leu Ala Asn Ile Phe Met Leu Ala Trp Pro
            340                 345                 350
Tyr Gly Tyr Pro Ala Leu Met Ser Ser Tyr Ala Phe Asn Arg Ser Thr
        355                 360                 365
Ala Tyr Asp Thr Ser Phe Gly Pro Pro His Asp Ser Gly Gly Ala Thr
    370                 375                 380
Arg Gly Pro Trp Asp Gly Gly Gly Ser Gln Pro Ala Cys Phe Asn Gln
385                 390                 395                 400
Ser Ile Gly Gly Trp Val Cys Glu His Arg Trp Arg Gly Ile Ala Asn
                405                 410                 415
Met Val Ala Phe Arg Asn Ala Thr Leu Pro Asn Trp Thr Val Thr Asp
            420                 425                 430
Trp Trp Asp Asn Gly Asn Asn Gln Ile Ala Phe Gly Arg Gly Asp Lys
        435                 440                 445
Gly Phe Val Val Ile Asn Arg Glu Asp Ala Ala Leu Thr Arg Asn Phe
    450                 455                 460
Lys Thr Ser Leu Pro Ala Gly Gln Tyr Cys Asp Val Ile Ser Gly Asp
465                 470                 475                 480
Phe Asn Asn Gly Gln Cys Thr Gly His Val Val Thr Val Asp Ala Gly
                485                 490                 495
```

```
Gly Tyr Val Thr Leu Thr Ala Gly Pro Asn Gly Ala Ala Ala Ile His
            500                 505                 510
Val Gly Ala Arg Leu Asp Gly Ala Ser Gln Pro Pro Thr Thr Ala Ser
        515                 520                 525
Val Thr Phe Asn Ala Ser Ala Asp Thr Phe Trp Gly Gln Asn Leu Phe
    530                 535                 540
Val Val Gly Asn His Ser Ala Leu Gly Asn Trp Ser Pro Ala Ala Ala
545                 550                 555                 560
Arg Pro Met Thr Trp Ile Ser Gly Ser Gly Thr Arg Gly Asn Trp Arg
                565                 570                 575
Ala Val Leu Asn Leu Pro Ala Asn Thr Thr Tyr Gln Tyr Lys Phe Ile
            580                 585                 590
Lys Lys Asp Gly Ala Gly Asn Val Val Trp Glu Gly Gly Gly Asn Arg
        595                 600                 605
Val Val Thr Thr Pro Ser Gly Gly Gly Ser Val Ser Thr Gly Gly Asn
    610                 615                 620
Trp Gln
625
```

<210> 133
<211> 1638
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 133

```
atgaataatg tgaaaaaagt atggttgtat tattctataa ttgctacctt agttatttcc      60
tttttcacac ctttttcaac agcacaagct aatactgcac ctgtcaacgg aacaatgatg     120
caatatttcg aatgggattt accgaatgat gggacgcttt ggacgaaagt aaaaaatgaa     180
gctaccaatc tttcttcgct aggtattaca gcgttatggc tccctccagc atataaagga     240
acgagccaaa gcgatgtcgg atatggcgtg tacgatttat atgaccttgg ggaatttaat     300
caaaaaggga cgatccgaac gaaatacgga acaaaagcac aatatattca agccatccaa     360
gctgccaaag ccgcagggat gcaagtatat gcagatgttg tatttaatca taaggcgggg     420
gctgacggca cagaatttgt cgatgcagtt gaggtaaacc cttctaatcg aaatcaagaa     480
acatctggca catatcaaat tcaagcatgg acaaaatttg attttcctgg tcgtggaaac     540
acatactcca gcttcaaatg gcgctggtat cattttgacg gtaccgattg ggatgaaagt     600
cgtaaattaa atcgtattta caaattccgc ggtacaggaa aagcgtggga ctgggaagtc     660
gatacagaaa acggaaacta tgattattta atgttcgctg atttagatat ggatcaccct     720
gaagttgtga cagagttaaa aaactgggga aaatggtatg taaatacgac aaatgtagac     780
ggatttcgtt tggatgccgt aaaacatatt aaatacagct ttttccctga ctggctaaca     840
tatgtacgta atcaaacagg aaaaaattta tttgctgttg gggaattttg gagctatgac     900
gtcaataagc tgcataacta cattacaaaa acaaatggat cgatgtcgtt atttgatgca     960
cctttgcata acaactttta tatcgcttcc aaatcgagtg gatattttga catgcgttat    1020
ttattgaata atacattaat gaaagatcaa ccttcactcg ctgtaacact tgtcgataac    1080
catgatacac aaccaggtca atctttacaa tcatgggtag aagcttggtt taaaccgctt    1140
gcttacgcct ttatttaac aagacaagag gggtatcctt gcgtatttta cggtgactat    1200
tacggaatcc cgaaatacaa tattccggga ttaaaaagta aaattgatcc gcttttaatt    1260
gctcgtcgtg attatgctta tggaacacaa cgtgattaca ttgatcatca agacattatc    1320
ggatggacac gagaaggcat tgatgcaaaa ccgaactctg gacttgcggc tttaattacc    1380
gacggccctg gcggaagtaa atggatgtat gtcggtaaaa aacatgctgg gaaagtgttt    1440
tatgatttaa ctggaaatcg aagtgacaca gtaacgatta atgcggacgg ttggggagaa    1500
tttaaagtaa acggcggctc cgtttcgatt tgggtggcta aaacatcaaa cgtcacattt    1560
acagtcaata acgccacaac aacaagtgga caaaacgtat atgttgttgg caacattcca    1620
gagctaggca attctttg                                                 1638
```

<210> 134
<211> 546
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 134

```
Met Asn Asn Val Lys Lys Val Trp Leu Tyr Tyr Ser Ile Ile Ala Thr
 1               5                  10                  15
```

```
Leu Val Ile Ser Phe Phe Thr Pro Phe Ser Thr Ala Gln Ala Asn Thr
            20                  25                  30
Ala Pro Val Asn Gly Thr Met Met Gln Tyr Phe Glu Trp Asp Leu Pro
        35                  40                  45
Asn Asp Gly Thr Leu Trp Thr Lys Val Lys Asn Glu Ala Thr Asn Leu
    50                  55                  60
Ser Ser Leu Gly Ile Thr Ala Leu Trp Leu Pro Pro Ala Tyr Lys Gly
65                  70                  75                  80
Thr Ser Gln Ser Asp Val Gly Tyr Gly Val Tyr Asp Leu Tyr Asp Leu
                85                  90                  95
Gly Glu Phe Asn Gln Lys Gly Thr Ile Arg Thr Lys Tyr Gly Thr Lys
            100                 105                 110
Ala Gln Tyr Ile Gln Ala Ile Gln Ala Ala Lys Ala Ala Gly Met Gln
        115                 120                 125
Val Tyr Ala Asp Val Val Phe Asn His Lys Ala Gly Ala Asp Gly Thr
    130                 135                 140
Glu Phe Val Asp Ala Val Glu Val Asn Pro Ser Asn Arg Asn Gln Glu
145                 150                 155                 160
Thr Ser Gly Thr Tyr Gln Ile Gln Ala Trp Thr Lys Phe Asp Phe Pro
                165                 170                 175
Gly Arg Gly Asn Thr Tyr Ser Ser Phe Lys Trp Arg Trp Tyr His Phe
            180                 185                 190
Asp Gly Thr Asp Trp Asp Glu Ser Arg Lys Leu Asn Arg Ile Tyr Lys
        195                 200                 205
Phe Arg Gly Thr Gly Lys Ala Trp Asp Trp Glu Val Asp Thr Glu Asn
    210                 215                 220
Gly Asn Tyr Asp Tyr Leu Met Phe Ala Asp Leu Asp Met Asp His Pro
225                 230                 235                 240
Glu Val Val Thr Glu Leu Lys Asn Trp Gly Lys Trp Tyr Val Asn Thr
                245                 250                 255
Thr Asn Val Asp Gly Phe Arg Leu Asp Ala Val Lys His Ile Lys Tyr
            260                 265                 270
Ser Phe Phe Pro Asp Trp Leu Thr Tyr Val Arg Asn Gln Thr Gly Lys
        275                 280                 285
Asn Leu Phe Ala Val Gly Glu Phe Trp Ser Tyr Asp Val Asn Lys Leu
    290                 295                 300
His Asn Tyr Ile Thr Lys Thr Asn Gly Ser Met Ser Leu Phe Asp Ala
305                 310                 315                 320
Pro Leu His Asn Asn Phe Tyr Ile Ala Ser Lys Ser Ser Gly Tyr Phe
                325                 330                 335
Asp Met Arg Tyr Leu Leu Asn Asn Thr Leu Met Lys Asp Gln Pro Ser
            340                 345                 350
Leu Ala Val Thr Leu Val Asp Asn His Asp Thr Gln Pro Gly Gln Ser
        355                 360                 365
Leu Gln Ser Trp Val Glu Ala Trp Phe Lys Pro Leu Ala Tyr Ala Phe
    370                 375                 380
Ile Leu Thr Arg Gln Glu Gly Tyr Pro Cys Val Phe Tyr Gly Asp Tyr
385                 390                 395                 400
Tyr Gly Ile Pro Lys Tyr Asn Ile Pro Gly Leu Lys Ser Lys Ile Asp
                405                 410                 415
Pro Leu Leu Ile Ala Arg Arg Asp Tyr Ala Tyr Gly Thr Gln Arg Asp
            420                 425                 430
Tyr Ile Asp His Gln Asp Ile Ile Gly Trp Thr Arg Glu Gly Ile Asp
        435                 440                 445
Ala Lys Pro Asn Ser Gly Leu Ala Ala Leu Ile Thr Asp Gly Pro Gly
    450                 455                 460
Gly Ser Lys Trp Met Tyr Val Gly Lys Lys His Ala Gly Lys Val Phe
465                 470                 475                 480
Tyr Asp Leu Thr Gly Asn Arg Ser Asp Thr Val Thr Ile Asn Ala Asp
                485                 490                 495
Gly Trp Gly Glu Phe Lys Val Asn Gly Gly Ser Val Ser Ile Trp Val
            500                 505                 510
Ala Lys Thr Ser Asn Val Thr Phe Thr Val Asn Asn Ala Thr Thr Thr
        515                 520                 525
Ser Gly Gln Asn Val Tyr Val Val Gly Asn Ile Pro Glu Leu Gly Asn
    530                 535                 540
Ser Leu
545
```

<210> 135
<211> 1935
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 135
gtgacaggca ccccgtcttt atacattcct ccacataaaa taaccataca gctttcaaat 60
ttgttgaaat gtataaaaat aaaaaatagt attgtaagcg ttaacatccg tcattataat 120
aacttcaaac gcgtttatgt tttaatgcaa acgtttgcat cctcatttta tttaaagaaa 180
ggatgtgtgt gcatgaatta tttgaaaaaa gtgtggttgt attacgctat cgtcgctacc 240
ttaatcattt cctttcttac gccctttca actgcacaag ccaacactgc accagtcaac 300
ggaacgatga tgcaatattt cgaatgggat ttaccgaatg atggcacact ttggacgaaa 360
gtaaaaaacg aagcaagcag cctttcttct ttaggtatta ctgcgttatg gttaccacct 420
gcatacaaag gaacgagcca aggggatgtc gggtatggcg tgtacgattt gtatgactta 480
ggagaattta atcaaaaagg gacgattcga acgaaatacg gaacaaaaac gcaatattta 540
caagccattc aagcggcaaa aagcgctggc atgcaagtat acgctgatgt cgtatttaat 600
cacaaggcgg gggcagatag tacagaatgg gttgacgcag tcgaagtgaa tccttctaat 660
cgaaaccaag aaacatctgg cacatatcaa attcaagcat ggacaaaatt tgatttccct 720
gaccgtggga acacatactc aagctttaaa tggcgctggt atcattttga cggtacggat 780
tgggatgaaa gtcgaaaact aaatcgcatt tacaaatttc gtggcacagg aaaagcatgg 840
gattgggaag tagacacaga gaacggaaac tatgactact taatgtttgc tgatttagat 900
atggatcacc ctgaagtcgt gacagagcta aaaaactggg gaacatggta cgtcaatacg 960
acaaatgtcg atgggtttcg cttagatgca gtaaagcata ttaaatatag ctttttccca 1020
gattggttaa catatgtgcg ctcacaaaca caaaaaaatc tgtttgcagt aggagaattt 1080
tggagctacg atgtcaataa actgcataac tacattacaa aaacaagtgg aaccatgtcg 1140
ttatttgatg cgccacttca taacaacttt tacactgctt caaaatctag cgggtatttt 1200
gacatgcgct atttgttaaa taatacgttg atgaaagacc agccttctct tgcggtcaca 1260
ctcgttgata atcatgacac gcaaccggga caatctttac aatcatgggt agagccttgg 1320
tttaagccgc ttgcttatgc ctttattttg acaagacaag aaggatatcc ttgcgtattt 1380
tacggcgact attacggcat ccctaaatac aatattccgg gattgaaaag taaaatcgat 1440
ccgcttctca ttgcccgtag agactacgca tacggaacac aacgtgatta tattgaccat 1500
caagacatta ttggatggac acgggaagga attgactcaa aaccgaactc tggacttgcg 1560
gctttaatta ctgacggtcc tggtggaagt aaatggatgt atgtaggtaa aaagcatgct 1620
ggaaaagtgt tttacgatct cactggaaat cgaagcgata cggtaacgat taatgcagac 1680
ggctggggag agtttaaagt aaacggtggc tccgtttcca tttgggttgc caaaacatca 1740
caagtcacgt ttaccgtcaa caatgcgaca acgacaagcg gacaaaatgt gtatgtcgtt 1800
ggcaacattc cagagctcgg aaattggaac acagcaaacg caatcaaaat gaccccatct 1860
tcttatccaa cgtggaaaac aaccattgct cttccacaag gaaaagcaat tggcggcgta 1920
cgccatggcc cttga 1935

<210> 136
<211> 644
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 136
Val Thr Gly Thr Pro Ser Leu Tyr Ile Pro Pro His Lys Ile Thr Ile
 1   5   10   15
Gln Leu Ser Asn Leu Leu Lys Cys Ile Lys Ile Lys Asn Ser Ile Val
            20   25   30
Ser Val Asn Ile Arg His Tyr Asn Asn Phe Lys Arg Val Tyr Val Leu
        35   40   45
Met Gln Thr Phe Ala Ser Ser Phe Tyr Leu Lys Lys Gly Cys Val Cys
    50   55   60
Met Asn Tyr Leu Lys Lys Val Trp Leu Tyr Tyr Ala Ile Val Ala Thr
65   70   75   80
Leu Ile Ile Ser Phe Leu Thr Pro Phe Ser Thr Ala Gln Ala Asn Thr
                85   90   95
Ala Pro Val Asn Gly Thr Met Met Gln Tyr Phe Glu Trp Asp Leu Pro
            100   105   110
Asn Asp Gly Thr Leu Trp Thr Lys Val Lys Asn Glu Ala Ser Ser Leu

```
        115                 120                 125
Ser Ser Leu Gly Ile Thr Ala Leu Trp Leu Pro Pro Ala Tyr Lys Gly
    130                 135                 140
Thr Ser Gln Gly Asp Val Gly Tyr Gly Val Tyr Asp Leu Tyr Asp Leu
145                 150                 155                 160
Gly Glu Phe Asn Gln Lys Gly Thr Ile Arg Thr Lys Tyr Gly Thr Lys
                165                 170                 175
Thr Gln Tyr Leu Gln Ala Ile Gln Ala Ala Lys Ser Ala Gly Met Gln
            180                 185                 190
Val Tyr Ala Asp Val Val Phe Asn His Lys Ala Gly Ala Asp Ser Thr
        195                 200                 205
Glu Trp Val Asp Ala Val Glu Val Asn Pro Ser Asn Arg Asn Gln Glu
    210                 215                 220
Thr Ser Gly Thr Tyr Gln Ile Gln Ala Trp Thr Lys Phe Asp Phe Pro
225                 230                 235                 240
Asp Arg Gly Asn Thr Tyr Ser Ser Phe Lys Trp Arg Trp Tyr His Phe
                245                 250                 255
Asp Gly Thr Asp Trp Asp Glu Ser Arg Lys Leu Asn Arg Ile Tyr Lys
            260                 265                 270
Phe Arg Gly Thr Gly Lys Ala Trp Asp Trp Glu Val Asp Thr Glu Asn
        275                 280                 285
Gly Asn Tyr Asp Tyr Leu Met Phe Ala Asp Leu Asp Met Asp His Pro
    290                 295                 300
Glu Val Val Thr Glu Leu Lys Asn Trp Gly Thr Trp Tyr Val Asn Thr
305                 310                 315                 320
Thr Asn Val Asp Gly Phe Arg Leu Asp Ala Val Lys His Ile Lys Tyr
                325                 330                 335
Ser Phe Phe Pro Asp Trp Leu Thr Tyr Val Arg Ser Gln Thr Gln Lys
            340                 345                 350
Asn Leu Phe Ala Val Gly Glu Phe Trp Ser Tyr Asp Val Asn Lys Leu
        355                 360                 365
His Asn Tyr Ile Thr Lys Thr Ser Gly Thr Met Ser Leu Phe Asp Ala
    370                 375                 380
Pro Leu His Asn Asn Phe Tyr Thr Ala Ser Lys Ser Ser Gly Tyr Phe
385                 390                 395                 400
Asp Met Arg Tyr Leu Leu Asn Asn Thr Leu Met Lys Asp Gln Pro Ser
                405                 410                 415
Leu Ala Val Thr Leu Val Asp Asn His Asp Thr Gln Pro Gly Gln Ser
            420                 425                 430
Leu Gln Ser Trp Val Glu Pro Trp Phe Lys Pro Leu Ala Tyr Ala Phe
        435                 440                 445
Ile Leu Thr Arg Gln Glu Gly Tyr Pro Cys Val Phe Tyr Gly Asp Tyr
    450                 455                 460
Tyr Gly Ile Pro Lys Tyr Asn Ile Pro Gly Leu Lys Ser Lys Ile Asp
465                 470                 475                 480
Pro Leu Leu Ile Ala Arg Arg Asp Tyr Ala Tyr Gly Thr Gln Arg Asp
                485                 490                 495
Tyr Ile Asp His Gln Asp Ile Ile Gly Trp Thr Arg Glu Gly Ile Asp
            500                 505                 510
Ser Lys Pro Asn Ser Gly Leu Ala Ala Leu Ile Thr Asp Gly Pro Gly
        515                 520                 525
Gly Ser Lys Trp Met Tyr Val Gly Lys Lys His Ala Gly Lys Val Phe
    530                 535                 540
Tyr Asp Leu Thr Gly Asn Arg Ser Asp Thr Val Thr Ile Asn Ala Asp
545                 550                 555                 560
Gly Trp Gly Glu Phe Lys Val Asn Gly Gly Ser Val Ser Ile Trp Val
                565                 570                 575
Ala Lys Thr Ser Gln Val Thr Phe Thr Val Asn Asn Ala Thr Thr Thr
            580                 585                 590
Ser Gly Gln Asn Val Tyr Val Val Gly Asn Ile Pro Glu Leu Gly Asn
        595                 600                 605
Trp Asn Thr Ala Asn Ala Ile Lys Met Thr Pro Ser Ser Tyr Pro Thr
    610                 615                 620
Trp Lys Thr Thr Ile Ala Leu Pro Gln Gly Lys Ala Ile Gly Gly Val
625                 630                 635                 640
Arg His Gly Pro
```

<210> 137
<211> 1320
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 137
gtgggacggg caggcttggc gcatcactcg aacacttccg ccaaggggac atacgggtca 60
cctctcgaac tgcgtccgga tcgccccgcc gtggccgggg cggtcgagct tgaagatgtc 120
cagcggggag ccgccgccga ggatcacccc ggcggcgtac tcgcccaggg cggggctcag 180
cttgaagccg tggccggagc cgcctcccag gagccagacg ttggaggccc gcggatggcg 240
gtcgaggagg aggtggccgt cggggctgtt ctcgtactgg cagacgcggg tctcgaccag 300
cggcgcgtcc ttcagggccg ggaaccggcg ggccacctcg gcccgggccg cttccagcag 360
ggccggggtg atcgtccgct cgcccgccgt gggatcgatg ggctcgcccc gggtgtcgtc 420
cgccaccttg aagccgcggt gctcgttgcc ggggatgccg tagtagatcc gctcgccgag 480
atcgacccag accggacagc cgccctcctg gaagcgcggg tcgcccggcg gcgtgccgaa 540
gaagaacacc tcctggcggg tgttgcggag gaaccgctca ccgatcacgt ccgggaacag 600
cccggccagc cagggaccgc aggcgaagac gtagaggtcg gccgcgagag tggagccgtc 660
cgaaaggtga agccgctcca agggccccgg gaccatggcg gcctgccggt actccccgcc 720
ctcgccctgg aacagctcca ccacggtccg gcaggcgcgc cgggcgaaca gggcgccggc 780
ttcctcctcg taccagatcg tgcggacgcc gtcgaaatcg acctgggggа agcggctccg 840
ggcctccccc tgagacagct cggcgaccgg cagccccgcg tcctccagaa aaggaaggga 900
gtcgcggacg tagctgtcgt cctcgccgca catccagagg accccggtcc ttttgtacag 960
ccggtaaccg gactggactt cggcgtcccg ccagagctcg aaggagcggg cgacccactc 1020
cacgtacaga cggtcgggtc cgtaggcgcc gcggatgatc cgcgtctcgc caccggagct 1080
ggagcgggag tgccccggac cccaggcgtc caggagggtc acccgggctc cgcggcggag 1140
gagatgcagg gcggtccagc cgccgaaggc gccggcgccg acgacggcga tatggggatg 1200
ggagggcatg gcgggcgtaa ggttatcgca gcccgatcct tcgctggcat cccatctccg 1260
accggagtat cctggaaaat tcgaagaagg agatcgacat gcaatcgaac ggaaacgtga 1320

<210> 138
<211> 439
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 138
Val Gly Arg Ala Gly Leu Ala His His Ser Asn Thr Ser Ala Lys Gly
 1   5   10   15
Thr Tyr Gly Ser Pro Leu Glu Leu Arg Pro Asp Arg Pro Ala Val Ala
            20   25   30
Gly Ala Val Glu Leu Glu Asp Val Gln Arg Gly Ala Ala Ala Glu Asp
        35   40   45
His Pro Gly Gly Val Leu Ala Gln Gly Gly Ala Gln Leu Glu Ala Val
    50   55   60
Ala Gly Ala Ala Ser Gln Glu Pro Asp Val Gly Gly Pro Arg Met Ala
65   70   75   80
Val Glu Glu Glu Val Ala Val Gly Ala Val Leu Val Leu Ala Asp Ala
                85   90   95
Gly Leu Asp Gln Arg Arg Val Leu Gln Gly Arg Glu Pro Ala Gly His
            100   105   110
Leu Gly Pro Gly Arg Phe Gln Gln Gly Arg Gly Asp Arg Pro Leu Ala
        115   120   125
Arg Arg Gly Ile Asp Gly Leu Ala Pro Gly Val Val Arg His Leu Glu
    130   135   140
Ala Ala Val Leu Val Ala Gly Asp Ala Val Val Asp Pro Leu Ala Glu
145   150   155   160
Ile Asp Pro Asp Arg Thr Ala Ala Leu Leu Glu Ala Arg Val Ala Arg
                165   170   175
Arg Arg Ala Glu Glu Glu His Leu Leu Ala Gly Val Ala Glu Glu Pro
            180   185   190
Leu Thr Asp His Val Arg Glu Gln Pro Gly Gln Pro Gly Thr Ala Gly
        195   200   205
Glu Asp Val Glu Val Gly Arg Glu Ser Gly Ala Val Arg Lys Val Lys

210 215 220
Pro Leu Gln Gly Pro Arg Asp His Gly Gly Leu Pro Val Leu Pro Ala
225 230 235 240
Leu Ala Leu Glu Gln Leu His His Gly Pro Ala Gly Ala Pro Gly Glu
245 250 255
Gln Gly Ala Gly Phe Leu Leu Val Pro Asp Arg Ala Asp Ala Val Glu
260 265 270
Ile Asp Leu Gly Glu Ala Ala Pro Gly Leu Pro Leu Arg Gln Leu Gly
275 280 285
Asp Arg Gln Pro Arg Val Leu Gln Lys Arg Lys Gly Val Ala Asp Val
290 295 300
Ala Val Val Leu Ala Ala His Pro Glu Asp Pro Gly Pro Phe Val Gln
305 310 315 320
Pro Val Thr Gly Leu Asp Phe Gly Val Pro Pro Glu Leu Glu Gly Ala
325 330 335
Gly Asp Pro Leu His Val Gln Thr Val Gly Ser Val Gly Ala Ala Asp
340 345 350
Asp Pro Arg Leu Ala Thr Gly Ala Gly Ala Gly Val Pro Arg Thr Pro
355 360 365
Gly Val Gln Glu Gly His Pro Gly Ser Ala Ala Glu Glu Met Gln Gly
370 375 380
Gly Pro Ala Ala Glu Gly Ala Gly Ala Asp Asp Gly Asp Met Gly Met
385 390 395 400
Gly Gly His Gly Gly Arg Lys Val Ile Ala Ala Arg Ser Phe Ala Gly
405 410 415
Ile Pro Ser Pro Thr Gly Val Ser Trp Lys Ile Arg Arg Arg Arg Ser
420 425 430
Thr Cys Asn Arg Thr Glu Thr
435

<210> 139
<211> 1524
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 139
atgaaaacat tcaaccttaa acccacactt ttacctttaa ctttgctgct gagttcgccg 60
gtattggcgg cacaaaatgg aactatgatg cagtatttcc attggtatgt gccaaatgac 120
ggcgcactct ggacacaagt tgaaaacaat gcgccagcac tatccgacaa cggttttaca 180
gcgctgtggt tgccaccagc atataaaggc gcaggtggta gcaacgacgt tggttacggt 240
gtttacgata tgtatgactt aggggagttt gatcaaaaag gatcggtacg aactaagtac 300
ggcaccaaag accaatatct aaatgccatc aaagcagcac acaaaaacaa tatccaaatt 360
tatggtgacg tagtgttcaa ccatcgtggc ggtgcagatg gcaagtcgtg ggtcgatacc 420
aagcgtgtgg attggaataa ccgcaatatt gaacttggcg ataaatggat tgaagcatgg 480
gttgaattta gcttcccagg acgtaacgat aaatactcag acttccattg gacgtggtat 540
cactttgatg gcgtcgattg ggatgacgca ggtaaagaga aagcgatctt taaattcaaa 600
ggtgatggta aagcatggga ttgggaagtc agttctgaaa aaggcaacta tgactacctc 660
atgtacgcag acttagacat ggatcaccca gaagtgaagc aagagctgaa agattggggt 720
gaatggtact taaacatgac gggtgttgat ggcttccgaa tggatgcagt gaaacacatc 780
aaatatcagt acctacaaga gtggatcgat tacttgcgta agaaaacggg caaagagctc 840
tttaccgttg gtgagtactg gaactacgac gtgaacaatc tgcacaactt tatgactaag 900
acttctggca gcatgtcatt gtttgatgcg cctttacata tgaacttcta taacgcttca 960
cgctctggtg gcaactttga tatgcgccga atcatggatg gcaccttgat gaaagacaac 1020
ccagtgaaag cagtaacact ggttgagaac catgatacgc aaccactaca ggccttagag 1080
tctccggtgg attggtggtt caaaccactt gcgtacgcgt tcattttgct tcgtgaggaa 1140
ggttatccgt cagtcttcta cgcagattac tacggtgcgc aatacagcga taaagggcac 1200
gatatcaaca tggtgaaagt gccttacatt gagcaattgg tgaaagcgcg taaagattat 1260
gcttatggta aacaacattc ttaccttgac cactgggatg tgattggttg gacacgagaa 1320
ggggatgcgg aacatccgaa ctctatggcg gttatcatga gtgatggtcc tggcggaaca 1380
aagtggatgt acacaggttc accgagcaca cgttatgtcg ataaactagg tattcgtacc 1440
gaagaagtat ggactaacgc tagtggatgg gccgaattcc cagtgaacgg cggatcggtt 1500
tctgtttggg ttggcgttaa ataa 1524

<210> 140
<211> 507

<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 140

```
Met Lys Thr Phe Asn Leu Lys Pro Thr Leu Leu Pro Leu Thr Leu Leu
 1               5                  10                  15
Leu Ser Ser Pro Val Leu Ala Ala Gln Asn Gly Thr Met Met Gln Tyr
            20                  25                  30
Phe His Trp Tyr Val Pro Asn Asp Gly Ala Leu Trp Thr Gln Val Glu
        35                  40                  45
Asn Asn Ala Pro Ala Leu Ser Asp Asn Gly Phe Thr Ala Leu Trp Leu
    50                  55                  60
Pro Pro Ala Tyr Lys Gly Ala Gly Gly Ser Asn Asp Val Gly Tyr Gly
65                  70                  75                  80
Val Tyr Asp Met Tyr Asp Leu Gly Glu Phe Asp Gln Lys Gly Ser Val
                85                  90                  95
Arg Thr Lys Tyr Gly Thr Lys Asp Gln Tyr Leu Asn Ala Ile Lys Ala
            100                 105                 110
Ala His Lys Asn Asn Ile Gln Ile Tyr Gly Asp Val Val Phe Asn His
        115                 120                 125
Arg Gly Gly Ala Asp Gly Lys Ser Trp Val Asp Thr Lys Arg Val Asp
    130                 135                 140
Trp Asn Asn Arg Asn Ile Glu Leu Gly Asp Lys Trp Ile Glu Ala Trp
145                 150                 155                 160
Val Glu Phe Ser Phe Pro Gly Arg Asn Asp Lys Tyr Ser Asp Phe His
                165                 170                 175
Trp Thr Trp Tyr His Phe Asp Gly Val Asp Trp Asp Asp Ala Gly Lys
            180                 185                 190
Glu Lys Ala Ile Phe Lys Phe Lys Gly Asp Gly Lys Ala Trp Asp Trp
        195                 200                 205
Glu Val Ser Ser Glu Lys Gly Asn Tyr Asp Tyr Leu Met Tyr Ala Asp
    210                 215                 220
Leu Asp Met Asp His Pro Glu Val Lys Gln Glu Leu Lys Asp Trp Gly
225                 230                 235                 240
Glu Trp Tyr Leu Asn Met Thr Gly Val Asp Gly Phe Arg Met Asp Ala
                245                 250                 255
Val Lys His Ile Lys Tyr Gln Tyr Leu Gln Glu Trp Ile Asp Tyr Leu
            260                 265                 270
Arg Lys Lys Thr Gly Lys Glu Leu Phe Thr Val Gly Glu Tyr Trp Asn
        275                 280                 285
Tyr Asp Val Asn Asn Leu His Asn Phe Met Thr Lys Thr Ser Gly Ser
    290                 295                 300
Met Ser Leu Phe Asp Ala Pro Leu His Met Asn Phe Tyr Asn Ala Ser
305                 310                 315                 320
Arg Ser Gly Gly Asn Phe Asp Met Arg Arg Ile Met Asp Gly Thr Leu
                325                 330                 335
Met Lys Asp Asn Pro Val Lys Ala Val Thr Leu Val Glu Asn His Asp
            340                 345                 350
Thr Gln Pro Leu Gln Ala Leu Glu Ser Pro Val Asp Trp Trp Phe Lys
        355                 360                 365
Pro Leu Ala Tyr Ala Phe Ile Leu Leu Arg Glu Glu Gly Tyr Pro Ser
    370                 375                 380
Val Phe Tyr Ala Asp Tyr Tyr Gly Ala Gln Tyr Ser Asp Lys Gly His
385                 390                 395                 400
Asp Ile Asn Met Val Lys Val Pro Tyr Ile Glu Gln Leu Val Lys Ala
                405                 410                 415
Arg Lys Asp Tyr Ala Tyr Gly Lys Gln His Ser Tyr Leu Asp His Trp
            420                 425                 430
Asp Val Ile Gly Trp Thr Arg Glu Gly Asp Ala Glu His Pro Asn Ser
        435                 440                 445
Met Ala Val Ile Met Ser Asp Gly Pro Gly Gly Thr Lys Trp Met Tyr
    450                 455                 460
Thr Gly Ser Pro Ser Thr Arg Tyr Val Asp Lys Leu Gly Ile Arg Thr
465                 470                 475                 480
Glu Glu Val Trp Thr Asn Ala Ser Gly Trp Ala Glu Phe Pro Val Asn
```

485 490 495
Gly Gly Ser Val Ser Val Trp Val Gly Val Lys
500 505

<210> 141
<211> 1401
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 141
atgaaaccaa taaataccct actcatatcc gcccttgctg tttgttcttt cagttccgcg 60
acttacgccg atactatttt gcacgcgttc aattggaagt attcagatgt gacggccaac 120
gcgaatcaaa ttgctcaagc tggttataag aaagtgcttg ttgcgcctgc aatgaaatcg 180
agtggcagcc aatggtgggc tcgctatcaa cctcaagatc tacgcactat cgattctcct 240
ttgggcaata aacaagattt agccgcaatg attgccgcac tcaaaggtgt gggcgtcgat 300
gtgtatgccg atgtggtact caaccatatg gcgaatgaaa gctggaagcg aagtgacttg 360
aattaccctg gcacagaagt gctaaacgat tatgctagcc gttcaagcta ctatgctgac 420
cagactctgt ttggcaacct agcacaaggt tatgtgtcag cgaacgactt tcatccagcg 480
ggctgtattt cagattggaa cgaccctggt catgttcagt attggcgttt gtgtggcgca 540
gatggtgatg taggtttacc tgaccttgat ccaaacaact gggtggtttc acaacagcgt 600
ttgtatctga aagcgctaaa agatatgggc atcaaagggt tccgaattga tgcagtgaag 660
cacatgagcc aataccaaat cgatcaggta ttcacgtctg aaattactgc gaacatgcat 720
gtgtttggtg aagtgattac tagcggtgga gcagggaata gcggctatga atcgttctta 780
gcgccttacc tgaataatac taatcactct gcctacgatt tcccgctgtt tgcatcgatt 840
cgctcggcat tttctatggg gggcggttta aatcaactgc atgatcctaa agcgtacggt 900
caggcacttg atgataatcg ctcgatcacc tttgcgatca cacatgatat tccaaccaat 960
gacggcttcc gctaccaaat tatggaccca caagacgagc agcttgctta cgcgtatatc 1020
cttggtaaag acggtggcac gccgctgatc tacagtgatg atcttcctga ttctgaagac 1080
aaggataacg gtcgttgggg caatgtttgg aacagttcga caatgaaaaa catgttgagc 1140
ttccataacg cgatgcaagg caaaacaatg acgatgattt ctagcgacca ttgcactttg 1200
ttgtttaagc gtggcaaaga aggtgttgtg ggtattaaca agtgtggtga aacgcgtggc 1260
gtgacggttg atacctacca acatgagttt aattggcatg ttcaatacaa agacgtgtta 1320
agcagcgcaa cagaaaccgt gacttctcgt taccatacgt tcaatctacc accacgcagt 1380
gcgcgtatgt ttaagctgta g 1401

<210> 142
<211> 466
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 142
Met Lys Pro Ile Asn Thr Leu Leu Ile Ser Ala Leu Ala Val Cys Ser
1 5 10 15
Phe Ser Ser Ala Thr Tyr Ala Asp Thr Ile Leu His Ala Phe Asn Trp
20 25 30
Lys Tyr Ser Asp Val Thr Ala Asn Ala Asn Gln Ile Ala Gln Ala Gly
35 40 45
Tyr Lys Lys Val Leu Val Ala Pro Ala Met Lys Ser Ser Gly Ser Gln
50 55 60
Trp Trp Ala Arg Tyr Gln Pro Gln Asp Leu Arg Thr Ile Asp Ser Pro
65 70 75 80
Leu Gly Asn Lys Gln Asp Leu Ala Ala Met Ile Ala Ala Leu Lys Gly
85 90 95
Val Gly Val Asp Val Tyr Ala Asp Val Val Leu Asn His Met Ala Asn
100 105 110
Glu Ser Trp Lys Arg Ser Asp Leu Asn Tyr Pro Gly Thr Glu Val Leu
115 120 125
Asn Asp Tyr Ala Ser Arg Ser Ser Tyr Tyr Ala Asp Gln Thr Leu Phe
130 135 140
Gly Asn Leu Ala Gln Gly Tyr Val Ser Ala Asn Asp Phe His Pro Ala
145 150 155 160
Gly Cys Ile Ser Asp Trp Asn Asp Pro Gly His Val Gln Tyr Trp Arg

```
                165                 170                 175
Leu Cys Gly Ala Asp Gly Asp Val Gly Leu Pro Asp Leu Asp Pro Asn
            180                 185                 190
Asn Trp Val Val Ser Gln Gln Arg Leu Tyr Leu Lys Ala Leu Lys Asp
        195                 200                 205
Met Gly Ile Lys Gly Phe Arg Ile Asp Ala Val Lys His Met Ser Gln
    210                 215                 220
Tyr Gln Ile Asp Gln Val Phe Thr Ser Glu Ile Thr Ala Asn Met His
225                 230                 235                 240
Val Phe Gly Glu Val Ile Thr Ser Gly Gly Ala Gly Asn Ser Gly Tyr
                245                 250                 255
Glu Ser Phe Leu Ala Pro Tyr Leu Asn Asn Thr Asn His Ser Ala Tyr
            260                 265                 270
Asp Phe Pro Leu Phe Ala Ser Ile Arg Ser Ala Phe Ser Met Gly Gly
        275                 280                 285
Gly Leu Asn Gln Leu His Asp Pro Lys Ala Tyr Gly Gln Ala Leu Asp
    290                 295                 300
Asp Asn Arg Ser Ile Thr Phe Ala Ile Thr His Asp Ile Pro Thr Asn
305                 310                 315                 320
Asp Gly Phe Arg Tyr Gln Ile Met Asp Pro Gln Asp Glu Gln Leu Ala
                325                 330                 335
Tyr Ala Tyr Ile Leu Gly Lys Asp Gly Gly Thr Pro Leu Ile Tyr Ser
            340                 345                 350
Asp Asp Leu Pro Asp Ser Glu Asp Lys Asp Asn Gly Arg Trp Gly Asn
        355                 360                 365
Val Trp Asn Ser Ser Thr Met Lys Asn Met Leu Ser Phe His Asn Ala
    370                 375                 380
Met Gln Gly Lys Thr Met Thr Met Ile Ser Ser Asp His Cys Thr Leu
385                 390                 395                 400
Leu Phe Lys Arg Gly Lys Glu Gly Val Val Gly Ile Asn Lys Cys Gly
                405                 410                 415
Glu Thr Arg Gly Val Thr Val Asp Thr Tyr Gln His Glu Phe Asn Trp
            420                 425                 430
His Val Gln Tyr Lys Asp Val Leu Ser Ser Ala Thr Glu Thr Val Thr
        435                 440                 445
Ser Arg Tyr His Thr Phe Asn Leu Pro Pro Arg Ser Ala Arg Met Phe
    450                 455                 460
Lys Leu
465
```

<210> 143
<211> 1422
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 143

```
atgccaaaga gcacttttac caaatccata acaaaatcac ttcttgctac ttccgttgtt      60
gtaagcttat tgcctgccta cgcacaggcc gacactatct tgcatgcctt taactggaaa     120
tacagcgaca ttacccgcca agcagagcaa attgcgcaag ctggttataa aaaagtactg     180
atttcaccgc cgctgaagtc cacaggccca caatggtggg cacgttacca accacaggac     240
attcgagtga ttgactcccc tgtcggcaac aagcaagatt tacaagccct cattgcagcc     300
ttaaaggcac aaggcgttga agtatacgca gacatcgtac tcaaccacat ggccaacgaa     360
agctggaaac gagacgatct gaactacccg ggaagtgatt tacttaccca atacagccaa     420
aatatggctt acatgaacca gcaaaaattg tttggagatt tagagcaaaa tcagttctct     480
gccaatgatt ttcacccggc tggctgcatt actgattgga gtaacccggg gcatgttcaa     540
tactggcgct tatgtggtgg taatggtgac actgggttac ctgatcttga tcctaactcg     600
tgggtgatcg atcaacaaaa acgttattta cgtgctttga aagacatggg aataaagggc     660
ttccgagttg atgcggtaaa acacatgagc gattaccaaa tcaaccaagt gtttacgcca     720
gacatcatcg caggcttaca tgtatttggt gaagtgatca ccagtggtgg caagggcagc     780
aatgactacc actcttttct ggaaccgtat ttaaataaca ccaatcacgc cgcgtatgac     840
ttcccgctat ttgcctctat ccgaaatgca tttagttatc atggcagctt gtctcaatta     900
catgatccac aagcttacgg gcaagcactt cctaacgaca gagccattac tttcaccatc     960
actcacgaca ttccaaccaa tgatggtttc cgttaccaaa tcatggatcc aaccagtgaa    1020
aaactcgcgt acgcgtacat tctaggcaaa gatgggggta gcccacttat ctatagcgat    1080
gctttagacc caagtgaaga taaagataag ggccgctggc gtgatgtatg gaaccaagaa    1140
```

tacatggtta acatgatcag cttccacaac aaggtgcaag gtaaaagcat ggaggtcatg 1200
tacagcgatc aatgcttgct ggtctttaaa cgtgaaaaac aaggcttagt cggtattaat 1260
aagtgcgctg aaagccgtac ctacaccata gatacccatc gttttgaatt taactggtac 1320
caaccgtaca acgacacatt aagccagcac agcgagacct ttagcagccg ttatcatgct 1380
ctgaccattc cggcgcaaac agcacgaatg ttggcgctat aa 1422

<210> 144
<211> 473
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 144
Met Pro Lys Ser Thr Phe Thr Lys Ser Ile Thr Lys Ser Leu Leu Ala
1 5 10 15
Thr Ser Val Val Val Ser Leu Leu Pro Ala Tyr Ala Gln Ala Asp Thr
20 25 30
Ile Leu His Ala Phe Asn Trp Lys Tyr Ser Asp Ile Thr Arg Gln Ala
35 40 45
Glu Gln Ile Ala Gln Ala Gly Tyr Lys Lys Val Leu Ile Ser Pro Pro
50 55 60
Leu Lys Ser Thr Gly Pro Gln Trp Trp Ala Arg Tyr Gln Pro Gln Asp
65 70 75 80
Ile Arg Val Ile Asp Ser Pro Val Gly Asn Lys Gln Asp Leu Gln Ala
85 90 95
Leu Ile Ala Ala Leu Lys Ala Gln Gly Val Glu Val Tyr Ala Asp Ile
100 105 110
Val Leu Asn His Met Ala Asn Glu Ser Trp Lys Arg Asp Asp Leu Asn
115 120 125
Tyr Pro Gly Ser Asp Leu Leu Thr Gln Tyr Ser Gln Asn Met Ala Tyr
130 135 140
Met Asn Gln Gln Lys Leu Phe Gly Asp Leu Glu Gln Asn Gln Phe Ser
145 150 155 160
Ala Asn Asp Phe His Pro Ala Gly Cys Ile Thr Asp Trp Ser Asn Pro
165 170 175
Gly His Val Gln Tyr Trp Arg Leu Cys Gly Gly Asn Gly Asp Thr Gly
180 185 190
Leu Pro Asp Leu Asp Pro Asn Ser Trp Val Ile Asp Gln Gln Lys Arg
195 200 205
Tyr Leu Arg Ala Leu Lys Asp Met Gly Ile Lys Gly Phe Arg Val Asp
210 215 220
Ala Val Lys His Met Ser Asp Tyr Gln Ile Asn Gln Val Phe Thr Pro
225 230 235 240
Asp Ile Ile Ala Gly Leu His Val Phe Gly Glu Val Ile Thr Ser Gly
245 250 255
Gly Lys Gly Ser Asn Asp Tyr His Ser Phe Leu Glu Pro Tyr Leu Asn
260 265 270
Asn Thr Asn His Ala Ala Tyr Asp Phe Pro Leu Phe Ala Ser Ile Arg
275 280 285
Asn Ala Phe Ser Tyr His Gly Ser Leu Ser Gln Leu His Asp Pro Gln
290 295 300
Ala Tyr Gly Gln Ala Leu Pro Asn Asp Arg Ala Ile Thr Phe Thr Ile
305 310 315 320
Thr His Asp Ile Pro Thr Asn Asp Gly Phe Arg Tyr Gln Ile Met Asp
325 330 335
Pro Thr Ser Glu Lys Leu Ala Tyr Ala Tyr Ile Leu Gly Lys Asp Gly
340 345 350
Gly Ser Pro Leu Ile Tyr Ser Asp Ala Leu Asp Pro Ser Glu Asp Lys
355 360 365
Asp Lys Gly Arg Trp Arg Asp Val Trp Asn Gln Glu Tyr Met Val Asn
370 375 380
Met Ile Ser Phe His Asn Lys Val Gln Gly Lys Ser Met Glu Val Met
385 390 395 400
Tyr Ser Asp Gln Cys Leu Leu Val Phe Lys Arg Glu Lys Gln Gly Leu
405 410 415
Val Gly Ile Asn Lys Cys Ala Glu Ser Arg Thr Tyr Thr Ile Asp Thr

```
            420                 425                 430
His Arg Phe Glu Phe Asn Trp Tyr Gln Pro Tyr Asn Asp Thr Leu Ser
        435                 440                 445
Gln His Ser Glu Thr Phe Ser Ser Arg Tyr His Ala Leu Thr Ile Pro
    450                 455                 460
Ala Gln Thr Ala Arg Met Leu Ala Leu
465                 470

<210> 145
<211> 1542
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 145
atgttgaaaa ggattacggt agtctgttta ttatttattt tgctttttcc taatatatat      60
gggaggaata aggcggaagc agcaacgata aataatggaa cattaatgca gtattttgag     120
tggtacgctc cgaatgatgg gaatcattgg aatcgtttgc gttatgatgc tgaaagttta     180
gctcataagg gaatcacatc tgtatggata ccacctgcat ataaagggac ttcgcaaaat     240
gatgtagggt atggggccta tgatttatac gatttagggg agttcaatca aaaaggaacg     300
gtgcggacga aatatgggac aaaggcacag ttgaaatctg caattgacgc tttacataag     360
caaaacatcg acgtatacgg tgatgtagtt atgaatcata aaggtggggc tgattatact     420
gaaaccgtaa cagctgttga ggtagaccgt aacaatcgaa atattgaagt atcaggtgat     480
tatgaaatta gtgcgtggac gggttttaac tttccagggc gcagagatgc ttattctaat     540
ttcaaatgga aatggtatca ttttgacgga acggattggg atgaaggaag gaaattaaac     600
cgaatttata aatttagggg tataggtaaa gcgtgggact gggaagtgtc tagcgaaaat     660
ggaaattatg attatttgat gtatgcagat cttgattttg atcatccaga tgttgcgaat     720
gaaatgaaaa gttggggaac gtggtatgcg aatgaattaa atttagatgg atttcgttta     780
gatgctgtta aacatattga tcatgaatat ttacgcgatt gggtaaatca tgtcagacag     840
caaacgggga aagaaatgtt tacggtggct gaatattggc aaaatgatat ccagacttta     900
aacaattatt tggcgaaagt caattataat caatctgtat ttgatgcacc gcttcattac     960
aattttcatt atgcttcaac aggaaatggg aattatgata tgagaaatat tttaaatgga    1020
acagtaatga aaaatcatcc tgcactcgca gttactctcg ttgagaatca tgattctcaa    1080
cctgggcaat cattggaatc tgtagtaagt ccgtggttta agccgctggc atatgcattt    1140
attttaactc gtgcagaggg ctatccttca gtttttatg gtgattacta tgggacaagc    1200
ggaaatagta gttatgaaat tccagcgtta aaagataaaa ttgatccaat tttgacggca    1260
cgaaaaaact ttgcatatgg tacgcagcgt gattatttag accatccaga tgtgattggc    1320
tggacaagag aaggagatag tgtacatgct aagtctggtt tagcggcatt aatctccgat    1380
ggaccaggag gatcaaagtg gatggatgtt ggaaagaata acgctgggga agtatggtac    1440
gatattacgg gtaatcaaac aaatactgta acaattaata aagatggatc ggggcaattc    1500
catgtaagtg gaggctctgt ttctatatat gttcaacagt aa                       1542

<210> 146
<211> 513
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 146
Met Leu Lys Arg Ile Thr Val Val Cys Leu Leu Phe Ile Leu Leu Phe
 1               5                  10                  15
Pro Asn Ile Tyr Gly Arg Asn Lys Ala Glu Ala Ala Thr Ile Asn Asn
            20                  25                  30
Gly Thr Leu Met Gln Tyr Phe Glu Trp Tyr Ala Pro Asn Asp Gly Asn
        35                  40                  45
His Trp Asn Arg Leu Arg Tyr Asp Ala Glu Ser Leu Ala His Lys Gly
    50                  55                  60
Ile Thr Ser Val Trp Ile Pro Pro Ala Tyr Lys Gly Thr Ser Gln Asn
65                  70                  75                  80
Asp Val Gly Tyr Gly Ala Tyr Asp Leu Tyr Asp Leu Gly Glu Phe Asn
                85                  90                  95
Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly Thr Lys Ala Gln Leu Lys
            100                 105                 110
Ser Ala Ile Asp Ala Leu His Lys Gln Asn Ile Asp Val Tyr Gly Asp
```

```
        115                 120                 125
Val Val Met Asn His Lys Gly Gly Ala Asp Tyr Thr Glu Thr Val Thr
    130                 135                 140
Ala Val Glu Val Asp Arg Asn Asn Arg Asn Ile Glu Val Ser Gly Asp
145                 150                 155                 160
Tyr Glu Ile Ser Ala Trp Thr Gly Phe Asn Phe Pro Gly Arg Arg Asp
                165                 170                 175
Ala Tyr Ser Asn Phe Lys Trp Lys Trp Tyr His Phe Asp Gly Thr Asp
            180                 185                 190
Trp Asp Glu Gly Arg Lys Leu Asn Arg Ile Tyr Lys Phe Arg Gly Ile
        195                 200                 205
Gly Lys Ala Trp Asp Trp Glu Val Ser Ser Glu Asn Gly Asn Tyr Asp
    210                 215                 220
Tyr Leu Met Tyr Ala Asp Leu Asp Phe Asp His Pro Asp Val Ala Asn
225                 230                 235                 240
Glu Met Lys Ser Trp Gly Thr Trp Tyr Ala Asn Glu Leu Asn Leu Asp
                245                 250                 255
Gly Phe Arg Leu Asp Ala Val Lys His Ile Asp His Glu Tyr Leu Arg
            260                 265                 270
Asp Trp Val Asn His Val Arg Gln Gln Thr Gly Lys Glu Met Phe Thr
        275                 280                 285
Val Ala Glu Tyr Trp Gln Asn Asp Ile Gln Thr Leu Asn Asn Tyr Leu
    290                 295                 300
Ala Lys Val Asn Tyr Asn Gln Ser Val Phe Asp Ala Pro Leu His Tyr
305                 310                 315                 320
Asn Phe His Tyr Ala Ser Thr Gly Asn Gly Asn Tyr Asp Met Arg Asn
                325                 330                 335
Ile Leu Asn Gly Thr Val Met Lys Asn His Pro Ala Leu Ala Val Thr
            340                 345                 350
Leu Val Glu Asn His Asp Ser Gln Pro Gly Gln Ser Leu Glu Ser Val
        355                 360                 365
Val Ser Pro Trp Phe Lys Pro Leu Ala Tyr Ala Phe Ile Leu Thr Arg
    370                 375                 380
Ala Glu Gly Tyr Pro Ser Val Phe Tyr Gly Asp Tyr Tyr Gly Thr Ser
385                 390                 395                 400
Gly Asn Ser Ser Tyr Glu Ile Pro Ala Leu Lys Asp Lys Ile Asp Pro
                405                 410                 415
Ile Leu Thr Ala Arg Lys Asn Phe Ala Tyr Gly Thr Gln Arg Asp Tyr
            420                 425                 430
Leu Asp His Pro Asp Val Ile Gly Trp Thr Arg Glu Gly Asp Ser Val
        435                 440                 445
His Ala Lys Ser Gly Leu Ala Ala Leu Ile Ser Asp Gly Pro Gly Gly
    450                 455                 460
Ser Lys Trp Met Asp Val Gly Lys Asn Asn Ala Gly Glu Val Trp Tyr
465                 470                 475                 480
Asp Ile Thr Gly Asn Gln Thr Asn Thr Val Thr Ile Asn Lys Asp Gly
                485                 490                 495
Ser Gly Gln Phe His Val Ser Gly Gly Ser Val Ser Ile Tyr Val Gln
            500                 505                 510
Gln
```

<210> 147
<211> 2343
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 147
atgagcttaa ataactttaa ggtaaaactg cttagttttg ctgtgtcttc tgctgtattg 60
tcactggctc caaatttagc caatgctgca aattttgaaa gtgagatggt gataatccat 120
ccgtttcagt ggacatatga caatatagca aaagagtgta cagagtacct tggtccagcc 180
ggatttgacg gtgtacagat ttcccagcca gcggaacata agcgggctga aggagtatgg 240
tgggccgtat atcagccggt taattataag aattttacaa ccatgaccgg taacgaggag 300
cagcttaagg caatgatcaa gacctgtaat gatgcaggtg ttaaggtgtt cgctgacgct 360
gttttcaacc aaaaggctac agacggtgta ggctggggcg gttcaacttg gagttataag 420

```
aactaccctg acggattctc cggatcagat ttccatggag actgttccat tgacaaaagc     480
tatactgatg caaataatgt cagaacctgt gcactctcag gtatgccgga cgttgccaca     540
gataactccg ctactcagga aaagattgca gattacctcg cttctttaat gaatatgggg     600
gtctatggtt tccgtattga cgctgcaaag cacatgggat acaacgatat caactccatt     660
ctttcaaaaa ctgcacagaa gactggaaga agacctcctg catatctgga agtaatcgga     720
gccggtaacg aagctgccga cattcagccg gacaagtata cctttattga gaatgcggtt     780
gtaactgact tcggttatgt ctgggatgca aatgagagtt tcggaaaggg taattacggt     840
aaggcactgg aactcagtac ctggctcggt gcaaattcag aaacattcgt aaacaatcat     900
gatgatgaat ggggcagatg ctcagccggt agctgctcaa tgaaaactca gaattatgct     960
gattataatc tggctcagtc ctggcttgct gtatggcctg taggtacagt aagacagata    1020
tattccggtt attcattccc tgtaaaagat aatgatcctt atcgcgtcag tgatgcaact    1080
catgatcagg gcgggcctct tggtgccgac cgctgtgaag gtggctggtt gtgtcagcac    1140
cgtgtgtcct tcgttctcaa ttccccaaga tttgcgagag ctaccagagg tactgctgta    1200
tcaaccaagg gatttgacaa tggtgctttg tggtttaaca gaggaagcaa aggtttttat    1260
gcacagaata ctaccaacag tcctataacc cagacattct ctgttgaagt acctgacgga    1320
aattactgtg atatcttagg aacatcagat cctaagagca atccatgcgg agcagacgtt    1380
gtcgtaagcg gcggtaaggc tacctttact attcctgcaa agacagctgt ggctatctgt    1440
acagactcag actggtgcgg caaggggttt gatccttgtg aaagtgatcc gaccggtgct    1500
gcctgtgttt gtaaggggga aaccaccgtt aatggtgtgt gcgtcagctg gtgtaatgcg    1560
cattcatcaa atgaggaatg cacctgtgta ttgaatccga atgatgccaa ctgtcaggct    1620
gatattgaac ctaccaaggg taaactctgt tacgccggta cttcaaacgg gtggaaacag    1680
gatcctttaa catataaccg taaaacaggt ttctggacta ttaatctgac tcttgacggt    1740
gcaggtgata ccagcggagc tcagcgcttc aaggttacag acggatgttc atggaccgga    1800
acagtttacg gttcttcagg tactgccgga aagttggatg taaatacatc atcaaccggc    1860
gatgaacctg tgtctcttgt tggtgattat gttctttcca ttaacgataa gaccatggaa    1920
tatacattca ccaaggcaga tgaagtaact aatcagccac cggttgcatc atttaccgcg    1980
acagttaacg gtctgaccgt ttcttttgcc aataattcat ccgaccctga gaatgatgaa    2040
ttaacctaca gctggaattt cggtaatggt aaaacatcat ccgagaaagc tcctagcata    2100
acctatgaag aatccggtaa gtatactgtt actttaaagg ttactgattc agctaataac    2160
actgatacat ttactaaaga tataactgta acagcacctt ctagtggcaa gtacttaaag    2220
gttgcagtca gaggttcgca tgataattac ggaactgatc tgttaaccaa gaacggttct    2280
gattggaccg gcgtctttga attctttgga tccactagtg tcgacctgca ggcgcgcgag    2340
ctc                                                                  2343

<210> 148
<211> 781
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 148
Met Ser Leu Asn Asn Phe Lys Val Lys Leu Leu Ser Phe Ala Val Ser
 1               5                  10                  15
Ser Ala Val Leu Ser Leu Ala Pro Asn Leu Ala Asn Ala Ala Asn Phe
            20                  25                  30
Glu Ser Glu Met Val Ile Ile His Pro Phe Gln Trp Thr Tyr Asp Asn
        35                  40                  45
Ile Ala Lys Glu Cys Thr Glu Tyr Leu Gly Pro Ala Gly Phe Asp Gly
    50                  55                  60
Val Gln Ile Ser Gln Pro Ala Glu His Lys Arg Ala Glu Gly Val Trp
65                  70                  75                  80
Trp Ala Val Tyr Gln Pro Val Asn Tyr Lys Asn Phe Thr Thr Met Thr
                85                  90                  95
Gly Asn Glu Glu Gln Leu Lys Ala Met Ile Lys Thr Cys Asn Asp Ala
            100                 105                 110
Gly Val Lys Val Phe Ala Asp Ala Val Phe Asn Gln Lys Ala Thr Asp
        115                 120                 125
Gly Val Gly Trp Gly Gly Ser Thr Trp Ser Tyr Lys Asn Tyr Pro Asp
    130                 135                 140
Gly Phe Ser Gly Ser Asp Phe His Gly Asp Cys Ser Ile Asp Lys Ser
145                 150                 155                 160
Tyr Thr Asp Ala Asn Asn Val Arg Thr Cys Ala Leu Ser Gly Met Pro
                165                 170                 175
Asp Val Ala Thr Asp Asn Ser Ala Thr Gln Glu Lys Ile Ala Asp Tyr
            180                 185                 190
Leu Ala Ser Leu Met Asn Met Gly Val Tyr Gly Phe Arg Ile Asp Ala
```

```
        195                 200                 205
Ala Lys His Met Gly Tyr Asn Asp Ile Asn Ser Ile Leu Ser Lys Thr
    210                 215                 220
Ala Gln Lys Thr Gly Arg Arg Pro Pro Ala Tyr Leu Glu Val Ile Gly
225                 230                 235                 240
Ala Gly Asn Glu Ala Ala Asp Ile Gln Pro Asp Lys Tyr Thr Phe Ile
                245                 250                 255
Glu Asn Ala Val Val Thr Asp Phe Gly Tyr Val Trp Asp Ala Asn Glu
            260                 265                 270
Ser Phe Gly Lys Gly Asn Tyr Gly Lys Ala Leu Glu Leu Ser Thr Trp
        275                 280                 285
Leu Gly Ala Asn Ser Glu Thr Phe Val Asn Asn His Asp Asp Glu Trp
    290                 295                 300
Gly Arg Cys Ser Ala Gly Ser Cys Ser Met Lys Thr Gln Asn Tyr Ala
305                 310                 315                 320
Asp Tyr Asn Leu Ala Gln Ser Trp Leu Ala Val Trp Pro Val Gly Thr
                325                 330                 335
Val Arg Gln Ile Tyr Ser Gly Tyr Ser Phe Pro Val Lys Asp Asn Asp
            340                 345                 350
Pro Tyr Arg Val Ser Asp Ala Thr His Asp Gln Gly Gly Pro Leu Gly
        355                 360                 365
Ala Asp Arg Cys Glu Gly Gly Trp Leu Cys Gln His Arg Val Ser Phe
    370                 375                 380
Val Leu Asn Ser Pro Arg Phe Ala Arg Ala Thr Arg Gly Thr Ala Val
385                 390                 395                 400
Ser Thr Lys Gly Phe Asp Asn Gly Ala Leu Trp Phe Asn Arg Gly Ser
                405                 410                 415
Lys Gly Phe Tyr Ala Gln Asn Thr Thr Asn Ser Pro Ile Thr Gln Thr
            420                 425                 430
Phe Ser Val Glu Val Pro Asp Gly Asn Tyr Cys Asp Ile Leu Gly Thr
        435                 440                 445
Ser Asp Pro Lys Ser Asn Pro Cys Gly Ala Asp Val Val Val Ser Gly
    450                 455                 460
Gly Lys Ala Thr Phe Thr Ile Pro Ala Lys Thr Ala Val Ala Ile Cys
465                 470                 475                 480
Thr Asp Ser Asp Trp Cys Gly Lys Gly Val Asp Pro Cys Glu Ser Asp
                485                 490                 495
Pro Thr Gly Ala Ala Cys Val Cys Lys Gly Glu Thr Thr Val Asn Gly
            500                 505                 510
Val Cys Val Ser Trp Cys Asn Ala His Ser Ser Asn Glu Glu Cys Thr
        515                 520                 525
Cys Val Leu Asn Pro Asn Asp Ala Asn Cys Gln Ala Asp Ile Glu Pro
    530                 535                 540
Thr Lys Gly Lys Leu Cys Tyr Ala Gly Thr Ser Asn Gly Trp Lys Gln
545                 550                 555                 560
Asp Pro Leu Thr Tyr Asn Arg Lys Thr Gly Phe Trp Thr Ile Asn Leu
                565                 570                 575
Thr Leu Asp Gly Ala Gly Asp Thr Ser Gly Ala Gln Arg Phe Lys Val
            580                 585                 590
Thr Asp Gly Cys Ser Trp Thr Gly Thr Val Tyr Gly Ser Ser Gly Thr
        595                 600                 605
Ala Gly Lys Leu Asp Val Asn Thr Ser Ser Thr Gly Asp Glu Pro Val
    610                 615                 620
Ser Leu Val Gly Asp Tyr Val Leu Ser Ile Asn Asp Lys Thr Met Glu
625                 630                 635                 640
Tyr Thr Phe Thr Lys Ala Asp Glu Val Thr Asn Gln Pro Pro Val Ala
                645                 650                 655
Ser Phe Thr Ala Thr Val Asn Gly Leu Thr Val Ser Phe Ala Asn Asn
            660                 665                 670
Ser Ser Asp Pro Glu Asn Asp Glu Leu Thr Tyr Ser Trp Asn Phe Gly
        675                 680                 685
Asn Gly Lys Thr Ser Ser Glu Lys Ala Pro Ser Ile Thr Tyr Glu Glu
    690                 695                 700
Ser Gly Lys Tyr Thr Val Thr Leu Lys Val Thr Asp Ser Ala Asn Asn
705                 710                 715                 720
Thr Asp Thr Phe Thr Lys Asp Ile Thr Val Thr Ala Pro Ser Ser Gly
                725                 730                 735
Lys Tyr Leu Lys Val Ala Val Arg Gly Ser His Asp Asn Tyr Gly Thr
```

740 745 750
Asp Leu Leu Thr Lys Asn Gly Ser Asp Trp Thr Gly Val Phe Glu Phe
755 760 765
Phe Gly Ser Thr Ser Val Asp Leu Gln Ala Arg Glu Leu
770 775 780

<210> 149
<211> 2502
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 149
atgatcttaa gtaattttaa ggtaaaactt cttagttttg ctgtgtcttc tgctgtactg 60
acactggctg caaatgtcgc caatgccaag aattatgaaa gtgaaatggt tattattcat 120
ccatttcagt ggacatatga caatatagca aaagaatgta ctgagtatct gggacctgcg 180
ggatttgacg gggtgcagat ttcccaggcg gctgagcata aagatgccgg tggtgcatgg 240
tggggtacct accagcctgt aaacttcaag agttttacta ccatggttgg taatgaagaa 300
cagcttagag caatgattaa aacctgtaac gaggcaggtg ttaaggtctt tgccgatgcc 360
gtgattaatc agaaagccgg cgacggtgta ggtataggtg gttcaacttt cggaaattat 420
aattatcctg acggatttac cagtgatgat tttcatcata ataactgcag tataggtaat 480
aattattcag atgcatgggt agtaagattc tgtgacctca gtggcatgcc ggatatagca 540
actgataacg acagtaccag aaataagatt gctgattact tcgccagcct tatgaatatg 600
ggggtatacg gattccgtat tgatgctgcc aagcacttta gctatgatga tatagacgct 660
attgtagaga aaacagcaac caaagcaggc aggagacctc ctgtctatat ggaggttatc 720
ggtaatccgg gtcaagaggc ggatgatatc cagccgaaca agtatacatg gattgataat 780
gccgttgtaa cagattttac ttatgctaat agcatgcata atatttttaa cggaagcggt 840
tatgccaagg ctttgaacat ggggctaggg catgttgatg ctgaaaatgc cgaagtcttt 900
ataagtaatc atgataatga atggggaaga aagtctgccg gttcctgctc aataagaacc 960
cagaataatc cggattacca tctggctcag tcctggctcg cagtttggcc tttaggcaag 1020
gttagacaga tttattctgc atatcagttc ccggtctttg aagatagttg tgagcgggtc 1080
agtcagcaag cccatgatca gggcggtcct atcggggcag cccgctgtga aggtggctgg 1140
ttgtgtcagc accgtgtacc gtttgtgctc aattctccta gatttgcaag agcaaccaga 1200
gggacagtcg ttactactaa aggttttgat gacggagctt tgtggtttaa cagaggaagc 1260
aagggcttct atgcccagaa tactaccggc agttctataa ctcatacatt ctcagttgaa 1320
ttacctgatg gaaattactg tgatatcctt ggagcaaccg atccgaagaa taatccttgc 1380
ggagcggatg tcactgtaag cggaggtaaa gcaaccttta ccattccggc aaagaccgcc 1440
gtagctatct gtactgatga aaagtggtgt ggcaaggggg ttgacccttg tgaaagcgat 1500
cctaccggtt ccgcctgtgt atgtaagggt gaaaccacag ttaacggcgt atgtgtaagc 1560
tggtgtaatg ctcactcatc taatgaagaa tgtgcctgtg tgctaaatcc taatgacgct 1620
gagtgtcagg ccgacattga gccgaccaag ggtaaactct gctatgtagg tacctccaac 1680
aagtggactc aggaaccttt aacctataat cgcaagaccg gtttctggac tctcaacgtt 1740
gaacttgacg gtaaggggga taccagcggg gcgcagcgct ttaaagttac cgacggctgt 1800
tcatggcagg gtactgttta cggttcatca ggagtagaag gcagacttga cgtaaatact 1860
tcagccaccg gagatgaacc ggtttcactg acaggtaaat atgttctttc cataaatgat 1920
aagaccatgg aatacacatt cattcctgca ggcagtggaa acaagcctcc ggttgcgtca 1980
tttactccga ctgttaaaga tctgactgta tcttttgtca ataattcatc cgaccctgag 2040
aatgatgaat taacctacag ctggaatttc ggtaacggta aaacctcatc tgaaaagaat 2100
ccgagtgtta catatgataa agccggtaaa tatactgttt cactcaaagt aaccgatact 2160
gcaaacaaca ctgataccaa aacactggaa atcgatttaa catctcctgt taacggaaaa 2220
tattccaagg ttgcagtcag aggttcacat gataactacg gaacaaatct gttaaccagg 2280
aatggttcag aatggaccgg tatctttgaa ttcagtaaga caaccaaatt caagcttgaa 2340
gctctgcctc ctgcagctga ccagtgtatc ttcctcggcg gtaatcgagg tgaggcattg 2400
actgcctccg gtggatttat atctcttcct gccggaaggt atactataaa gtttaatgag 2460
gaaagcaagg ttcttactgc aggcgatgtt gactgcaccg gg 2502

<210> 150
<211> 834
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 150
Met Ile Leu Ser Asn Phe Lys Val Lys Leu Leu Ser Phe Ala Val Ser

```
 1               5                  10                  15
Ser Ala Val Leu Thr Leu Ala Ala Asn Val Ala Asn Ala Lys Asn Tyr
            20                  25                  30
Glu Ser Glu Met Val Ile Ile His Pro Phe Gln Trp Thr Tyr Asp Asn
        35                  40                  45
Ile Ala Lys Glu Cys Thr Glu Tyr Leu Gly Pro Ala Gly Phe Asp Gly
    50                  55                  60
Val Gln Ile Ser Gln Ala Ala Glu His Lys Asp Ala Gly Gly Ala Trp
65                  70                  75                  80
Trp Gly Thr Tyr Gln Pro Val Asn Phe Lys Ser Phe Thr Thr Met Val
                85                  90                  95
Gly Asn Glu Glu Gln Leu Arg Ala Met Ile Lys Thr Cys Asn Glu Ala
            100                 105                 110
Gly Val Lys Val Phe Ala Asp Ala Val Ile Asn Gln Lys Ala Gly Asp
        115                 120                 125
Gly Val Gly Ile Gly Gly Ser Thr Phe Gly Asn Tyr Asn Tyr Pro Asp
    130                 135                 140
Gly Phe Thr Ser Asp Asp Phe His His Asn Asn Cys Ser Ile Gly Asn
145                 150                 155                 160
Asn Tyr Ser Asp Ala Trp Val Val Arg Phe Cys Asp Leu Ser Gly Met
                165                 170                 175
Pro Asp Ile Ala Thr Asp Asn Asp Ser Thr Arg Asn Lys Ile Ala Asp
            180                 185                 190
Tyr Phe Ala Ser Leu Met Asn Met Gly Val Tyr Gly Phe Arg Ile Asp
        195                 200                 205
Ala Ala Lys His Phe Ser Tyr Asp Asp Ile Asp Ala Ile Val Glu Lys
    210                 215                 220
Thr Ala Thr Lys Ala Gly Arg Arg Pro Pro Val Tyr Met Glu Val Ile
225                 230                 235                 240
Gly Asn Pro Gly Gln Glu Ala Asp Asp Ile Gln Pro Asn Lys Tyr Thr
                245                 250                 255
Trp Ile Asp Asn Ala Val Val Thr Asp Phe Thr Tyr Ala Asn Ser Met
            260                 265                 270
His Asn Ile Phe Asn Gly Ser Gly Tyr Ala Lys Ala Leu Asn Met Gly
        275                 280                 285
Leu Gly His Val Asp Ala Glu Asn Ala Glu Val Phe Ile Ser Asn His
    290                 295                 300
Asp Asn Glu Trp Gly Arg Lys Ser Ala Gly Ser Cys Ser Ile Arg Thr
305                 310                 315                 320
Gln Asn Asn Pro Asp Tyr His Leu Ala Gln Ser Trp Leu Ala Val Trp
                325                 330                 335
Pro Leu Gly Lys Val Arg Gln Ile Tyr Ser Ala Tyr Gln Phe Pro Val
            340                 345                 350
Phe Glu Asp Ser Cys Glu Arg Val Ser Gln Gln Ala His Asp Gln Gly
        355                 360                 365
Gly Pro Ile Gly Ala Ala Arg Cys Glu Gly Gly Trp Leu Cys Gln His
    370                 375                 380
Arg Val Pro Phe Val Leu Asn Ser Pro Arg Phe Ala Arg Ala Thr Arg
385                 390                 395                 400
Gly Thr Val Val Thr Thr Lys Gly Phe Asp Asp Gly Ala Leu Trp Phe
                405                 410                 415
Asn Arg Gly Ser Lys Gly Phe Tyr Ala Gln Asn Thr Thr Gly Ser Ser
            420                 425                 430
Ile Thr His Thr Phe Ser Val Glu Leu Pro Asp Gly Asn Tyr Cys Asp
        435                 440                 445
Ile Leu Gly Ala Thr Asp Pro Lys Asn Asn Pro Cys Gly Ala Asp Val
    450                 455                 460
Thr Val Ser Gly Gly Lys Ala Thr Phe Thr Ile Pro Ala Lys Thr Ala
465                 470                 475                 480
Val Ala Ile Cys Thr Asp Glu Lys Trp Cys Gly Lys Gly Val Asp Pro
                485                 490                 495
Cys Glu Ser Asp Pro Thr Gly Ser Ala Cys Val Cys Lys Gly Glu Thr
            500                 505                 510
Thr Val Asn Gly Val Cys Val Ser Trp Cys Asn Ala His Ser Ser Asn
        515                 520                 525
Glu Glu Cys Ala Cys Val Leu Asn Pro Asn Asp Ala Glu Cys Gln Ala
    530                 535                 540
Asp Ile Glu Pro Thr Lys Gly Lys Leu Cys Tyr Val Gly Thr Ser Asn
```

```
545                 550                 555                 560
Lys Trp Thr Gln Glu Pro Leu Thr Tyr Asn Arg Lys Thr Gly Phe Trp
                565                 570                 575
Thr Leu Asn Val Glu Leu Asp Gly Lys Gly Asp Thr Ser Gly Ala Gln
            580                 585                 590
Arg Phe Lys Val Thr Asp Gly Cys Ser Trp Gln Gly Thr Val Tyr Gly
        595                 600                 605
Ser Ser Gly Val Glu Gly Arg Leu Asp Val Asn Thr Ser Ala Thr Gly
    610                 615                 620
Asp Glu Pro Val Ser Leu Thr Gly Lys Tyr Val Leu Ser Ile Asn Asp
625                 630                 635                 640
Lys Thr Met Glu Tyr Thr Phe Ile Pro Ala Gly Ser Gly Asn Lys Pro
                645                 650                 655
Pro Val Ala Ser Phe Thr Pro Thr Val Lys Asp Leu Thr Val Ser Phe
            660                 665                 670
Val Asn Asn Ser Ser Asp Pro Glu Asn Asp Glu Leu Thr Tyr Ser Trp
        675                 680                 685
Asn Phe Gly Asn Gly Lys Thr Ser Ser Glu Lys Asn Pro Ser Val Thr
    690                 695                 700
Tyr Asp Lys Ala Gly Lys Tyr Thr Val Ser Leu Lys Val Thr Asp Thr
705                 710                 715                 720
Ala Asn Asn Thr Asp Thr Lys Thr Leu Glu Ile Asp Leu Thr Ser Pro
                725                 730                 735
Val Asn Gly Lys Tyr Ser Lys Val Ala Val Arg Gly Ser His Asp Asn
            740                 745                 750
Tyr Gly Thr Asn Leu Leu Thr Arg Asn Gly Ser Glu Trp Thr Gly Ile
        755                 760                 765
Phe Glu Phe Ser Lys Thr Thr Lys Phe Lys Leu Glu Ala Leu Pro Pro
    770                 775                 780
Ala Ala Asp Gln Cys Ile Phe Leu Gly Gly Asn Arg Gly Glu Ala Leu
785                 790                 795                 800
Thr Ala Ser Gly Gly Phe Ile Ser Leu Pro Ala Gly Arg Tyr Thr Ile
                805                 810                 815
Lys Phe Asn Glu Glu Ser Lys Val Leu Thr Ala Gly Asp Val Asp Cys
            820                 825                 830
Thr Gly
```

<210> 151
<211> 2112
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 151

```
atgaaaacta ttctttcaac aatcatggtg atggcggctg cggctgccac caccgtagag      60
gctcaaggct ggccggaaaa ctacggcggc gtcatgttgc agggattcta ctgggattcc     120
tattcagcca ccaagtggac taaactggaa gcacaggctg acgagatctg caactatttc     180
tcgctggtat gggtaccaca gtcggcctat accggcagca gtacctccat gggctacgac     240
ccgctgtatt acttcgacca gcattcatcg ttcggcaccg aagagcagct acggtcgttc     300
atcagtacct acaagcagaa aggaactggc atcatagccg atgtagttgt caatcaccga     360
aagaatgtct caaactgggt ggatttcccg gccgagacct acaacggtgt aacctatcag     420
atggtaagca ccgacatcgt ttcgaacgat gacggcggaa aaacagccac ttgggcaaat     480
caaaacggct acagtctctc ctccaatgcc gacgaaggcg aaggctggga cggcatgcgc     540
gacctggacc acaagtcgca gaacgtgcag aaatcggttc ttgcctacac caaatatctg     600
gttgacgact taggctatac cggattccgc tacgatatgg taaagggatt tgacggatcg     660
catgtagccg actacaacac caatgccggc gtgcagttct ctgtcggcga atattgggac     720
ggcactgcat cgaaagttta cagttggatc aacagcacca aaaagagcga tgtgccgcag     780
tcggcagcct tcgacttcgc tttccgatac acctgccgcg atgccgtcaa caacaagaac     840
tgggcgaacc tgaagaacac ttccggtatc agcgatgccg attacaggcg ctattcggtt     900
acgtttgttg aaaatcacga tacggaatac cgttcagcta cggcttccca ggatcccatc     960
aagggtgata cggttgccct caatgcctgg atgctggcta tgccgggcac accttgtgtt    1020
ttcctgaaac attggaccga ctgcaaggaa gagatcaaga atctcatcga ggcacgtcgc    1080
ctggtcggta ttcacaacca gagcacctat gccgaatgga tgagcggtgc agcctacatc    1140
ggacgtaccg taacaggtac gaacggcacc ttacgtgttc tgtgcggctc ttatcagtat    1200
aatgtagccg ccaactacat tcagattctc tcaggcaaaa actataaata ctacgtactc    1260
```

aacacgctcg aggctccctg gatcgggaaa ggttccggct cgtacaccga aggtgaaacc 1320
gtaacggttc cgctcatcgc catatcggcc gatgccaatg ccaagctggt atataccacc 1380
gacggcacag accccaccgc aacctcaaca gccgtaacca gcggaacgga actgaccatc 1440
acttcggacg ccgtcctgaa ggttggtctg ctttccggcg gcatcgtcag gaacatacag 1500
agccgtacat tcaccttcca ggctgcaaac acctccgagt attacacagc caccatgcac 1560
gtatgcaacc agtccggagc tctcaatccg ctgtttgcct atgtttgggc aggaccggac 1620
aacgagcaga ttaacggcaa ctggccgggc accaagctca ccgctaccat taccgaaaac 1680
aaccttacct ggtacacgca gtcgttccag attccgaaga acgtggacta tgtcgtgaac 1740
tttgttttca ccacaaccgg cggcggtacg cagacagtgg atgttaccgg catgaaggcc 1800
gatgtctggt acattattaa cagtaccaag agcggcaaca agtacacggt aaccgacgtt 1860
acctcacagt attcttcgtt agaggccatc tttgatgaag aaaactccgg ctccttccct 1920
gtctatgacc tgcagggacg ccgcgtcagc gaaattagaa acaggacaat tatatcttca 1980
gaacggaaag aagatactca tcagataaac agaggttccg aaccattctc ctattatgaa 2040
aatcagacac ttagtaatct cagcactgct gggtttgggg ggcttgtaca ccatcagctg 2100
ctcctcgtcg gg 2112

<210> 152
<211> 704
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 152
Met Lys Thr Ile Leu Ser Thr Ile Met Val Met Ala Ala Ala Ala Ala
 1   5   10   15
Thr Thr Val Glu Ala Gln Gly Trp Pro Glu Asn Tyr Gly Gly Val Met
            20   25   30
Leu Gln Gly Phe Tyr Trp Asp Ser Tyr Ser Ala Thr Lys Trp Thr Lys
        35   40   45
Leu Glu Ala Gln Ala Asp Glu Ile Cys Asn Tyr Phe Ser Leu Val Trp
    50   55   60
Val Pro Gln Ser Ala Tyr Thr Gly Ser Ser Thr Ser Met Gly Tyr Asp
65   70   75   80
Pro Leu Tyr Tyr Phe Asp Gln His Ser Ser Phe Gly Thr Glu Glu Gln
                85   90   95
Leu Arg Ser Phe Ile Ser Thr Tyr Lys Gln Lys Gly Thr Gly Ile Ile
            100   105   110
Ala Asp Val Val Val Asn His Arg Lys Asn Val Ser Asn Trp Val Asp
        115   120   125
Phe Pro Ala Glu Thr Tyr Asn Gly Val Thr Tyr Gln Met Val Ser Thr
    130   135   140
Asp Ile Val Ser Asn Asp Asp Gly Gly Lys Thr Ala Thr Trp Ala Asn
145   150   155   160
Gln Asn Gly Tyr Ser Leu Ser Ser Asn Ala Asp Glu Gly Glu Gly Trp
                165   170   175
Asp Gly Met Arg Asp Leu Asp His Lys Ser Gln Asn Val Gln Lys Ser
            180   185   190
Val Leu Ala Tyr Thr Lys Tyr Leu Val Asp Asp Leu Gly Tyr Thr Gly
        195   200   205
Phe Arg Tyr Asp Met Val Lys Gly Phe Asp Gly Ser His Val Ala Asp
    210   215   220
Tyr Asn Thr Asn Ala Gly Val Gln Phe Ser Val Gly Glu Tyr Trp Asp
225   230   235   240
Gly Thr Ala Ser Lys Val Tyr Ser Trp Ile Asn Ser Thr Lys Lys Ser
                245   250   255
Asp Val Pro Gln Ser Ala Ala Phe Asp Phe Ala Phe Arg Tyr Thr Cys
            260   265   270
Arg Asp Ala Val Asn Asn Lys Asn Trp Ala Asn Leu Lys Asn Thr Ser
        275   280   285
Gly Ile Ser Asp Ala Asp Tyr Arg Arg Tyr Ser Val Thr Phe Val Glu
    290   295   300
Asn His Asp Thr Glu Tyr Arg Ser Ala Thr Ala Ser Gln Asp Pro Ile
305   310   315   320
Lys Gly Asp Thr Val Ala Leu Asn Ala Trp Met Leu Ala Met Pro Gly
                325   330   335
Thr Pro Cys Val Phe Leu Lys His Trp Thr Asp Cys Lys Glu Glu Ile

```
            340                 345                 350
Lys Asn Leu Ile Glu Ala Arg Arg Leu Val Gly Ile His Asn Gln Ser
        355                 360                 365
Thr Tyr Ala Glu Trp Met Ser Gly Ala Ala Tyr Ile Gly Arg Thr Val
    370                 375                 380
Thr Gly Thr Asn Gly Thr Leu Arg Val Leu Cys Gly Ser Tyr Gln Tyr
385                 390                 395                 400
Asn Val Ala Ala Asn Tyr Ile Gln Ile Leu Ser Gly Lys Asn Tyr Lys
                405                 410                 415
Tyr Tyr Val Leu Asn Thr Leu Glu Ala Pro Trp Ile Gly Lys Gly Ser
            420                 425                 430
Gly Ser Tyr Thr Glu Gly Glu Thr Val Thr Val Pro Leu Ile Ala Ile
        435                 440                 445
Ser Ala Asp Ala Asn Ala Lys Leu Val Tyr Thr Thr Asp Gly Thr Asp
    450                 455                 460
Pro Thr Ala Thr Ser Thr Ala Val Thr Ser Gly Thr Glu Leu Thr Ile
465                 470                 475                 480
Thr Ser Asp Ala Val Leu Lys Val Gly Leu Leu Ser Gly Gly Ile Val
                485                 490                 495
Arg Asn Ile Gln Ser Arg Thr Phe Thr Phe Gln Ala Ala Asn Thr Ser
            500                 505                 510
Glu Tyr Tyr Thr Ala Thr Met His Val Cys Asn Gln Ser Gly Ala Leu
        515                 520                 525
Asn Pro Leu Phe Ala Tyr Val Trp Ala Gly Pro Asp Asn Glu Gln Ile
    530                 535                 540
Asn Gly Asn Trp Pro Gly Thr Lys Leu Thr Ala Thr Ile Thr Glu Asn
545                 550                 555                 560
Asn Leu Thr Trp Tyr Thr Gln Ser Phe Gln Ile Pro Lys Asn Val Asp
                565                 570                 575
Tyr Val Val Asn Phe Val Phe Thr Thr Thr Gly Gly Gly Thr Gln Thr
            580                 585                 590
Val Asp Val Thr Gly Met Lys Ala Asp Val Trp Tyr Ile Ile Asn Ser
        595                 600                 605
Thr Lys Ser Gly Asn Lys Tyr Thr Val Thr Asp Val Thr Ser Gln Tyr
    610                 615                 620
Ser Ser Leu Glu Ala Ile Phe Asp Glu Glu Asn Ser Gly Ser Phe Pro
625                 630                 635                 640
Val Tyr Asp Leu Gln Gly Arg Arg Val Ser Glu Ile Arg Asn Arg Thr
                645                 650                 655
Ile Ile Ser Ser Glu Arg Lys Glu Asp Thr His Gln Ile Asn Arg Gly
            660                 665                 670
Ser Glu Pro Phe Ser Tyr Tyr Glu Asn Gln Thr Leu Ser Asn Leu Ser
        675                 680                 685
Thr Ala Gly Phe Gly Gly Leu Val His His Gln Leu Leu Leu Val Gly
    690                 695                 700
```

<210> 153
<211> 1620
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 153
atgccttcaa ttaatgcaag cgattgcaaa aaaaagggag ataggagtat gaagaggaaa 60
aaatggactg cgttagcact atctttacca ctagttatga gcttatcaac aaacatacaa 120
gcagaaacat tacataataa taagggtcaa aaggcgcaaa caggaaataa agacggaatt 180
ttttatgaac tgtatgttaa ttctttttat gatactgata gcaatggaca tggtgattta 240
aaaggcgtca caaagaaact tgattattta aatgatggaa atccaagaac aaataatgat 300
cttcaaataa acggtatctg gatgatgcct attaacacct ctcctagtta tcacaaatat 360
gatgtaacag attactataa tatcgatcct cagtatggaa gtttacaaga tttccgtgaa 420
ctaacaacag aagcgcataa acgcaacgta aaggtagtaa tagatcttgt tattaatcat 480
acaagcagtg agcatccttg gtttgtcgat gcattaaaaa ataaaaacag taagtatcga 540
gattactata tttgggctga taaaaataca gacttaaatg aaaaaggccc atggggtcaa 600
caagtatggc acaaagcgtc gaacggagag tatttctacg caacgttctg ggaagggatg 660
ccggacttaa actatgacaa ccctaaagta agagaagaaa tgattaaaat cgggaaattt 720
tggctcaaac aaggagctga tggctttcgt ctagatgcag ccatgcacat ctttaaaggg 780

caaacacctg aaggagcaaa gaaaaatatt gaatggtgga atgaattccg cgacgcgatg 840
agagaaacga atccaaatac gtatctagtt ggtgaaatat gggatcaacc agaagtagtt 900
gctccgtatt atcaatcgtt agattctaca tttaacttcg acttagcata taaaatcgtt 960
aattccgtta aaaatggtac tgatcaaggg gtagccgcgg cagctgttgc aacggatgag 1020
ttatataaaa catataatcc aaataaaatt gatggaacgt ttttaacgaa tcatgaccaa 1080
aatcgtgtaa tgagtgagtt aaatggtgat gtaaacaaag caaaatcagc agcctctatt 1140
ctgttgacac tccctggtaa tccgttcatt tattatggcg aagaaatcgg catgacaggc 1200
caaaaaccag atgagttgat tcgtgagcct ttccgttggt atgaagatga taaagaaggt 1260
caaacgagct gggagactcc agtatataac attgatcata atggtgtttc agttgaagca 1320
caagataaac aaaaagcttc tcttctaagc cattatcgta aaatgattcg tgttcgtcag 1380
caacacgatg aacttgtcaa aggtaattta gaacctattt ctgtcaataa ttcacaggtt 1440
gttgcctata atcgtacgta taaaaataaa tcaattcaag tgtaccataa tatttcagac 1500
aagccggtta cattaactgt ttcaaacaaa ggaaaactga ttttttctag tgaattagga 1560
gcaaaaaagg aaaaatcaac attagtaatt ccagcgaata cgacagtgct agtaaagtaa 1620

<210> 154
<211> 539
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 154
Met Pro Ser Ile Asn Ala Ser Asp Cys Lys Lys Lys Gly Asp Arg Ser
 1   5   10   15
Met Lys Arg Lys Lys Trp Thr Ala Leu Ala Leu Ser Leu Pro Leu Val
            20   25   30
Met Ser Leu Ser Thr Asn Ile Gln Ala Glu Thr Leu His Asn Asn Lys
        35   40   45
Gly Gln Lys Ala Gln Thr Gly Asn Lys Asp Gly Ile Phe Tyr Glu Leu
    50   55   60
Tyr Val Asn Ser Phe Tyr Asp Thr Asp Ser Asn Gly His Gly Asp Leu
65   70   75   80
Lys Gly Val Thr Lys Lys Leu Asp Tyr Leu Asn Asp Gly Asn Pro Arg
                85   90   95
Thr Asn Asn Asp Leu Gln Ile Asn Gly Ile Trp Met Met Pro Ile Asn
            100   105   110
Thr Ser Pro Ser Tyr His Lys Tyr Asp Val Thr Asp Tyr Tyr Asn Ile
        115   120   125
Asp Pro Gln Tyr Gly Ser Leu Gln Asp Phe Arg Glu Leu Thr Thr Glu
    130   135   140
Ala His Lys Arg Asn Val Lys Val Val Ile Asp Leu Val Ile Asn His
145   150   155   160
Thr Ser Ser Glu His Pro Trp Phe Val Asp Ala Leu Lys Asn Lys Asn
                165   170   175
Ser Lys Tyr Arg Asp Tyr Tyr Ile Trp Ala Asp Lys Asn Thr Asp Leu
            180   185   190
Asn Glu Lys Gly Pro Trp Gly Gln Gln Val Trp His Lys Ala Ser Asn
        195   200   205
Gly Glu Tyr Phe Tyr Ala Thr Phe Trp Glu Gly Met Pro Asp Leu Asn
    210   215   220
Tyr Asp Asn Pro Lys Val Arg Glu Glu Met Ile Lys Ile Gly Lys Phe
225   230   235   240
Trp Leu Lys Gln Gly Ala Asp Gly Phe Arg Leu Asp Ala Ala Met His
                245   250   255
Ile Phe Lys Gly Gln Thr Pro Glu Gly Ala Lys Lys Asn Ile Glu Trp
            260   265   270
Trp Asn Glu Phe Arg Asp Ala Met Arg Glu Thr Asn Pro Asn Thr Tyr
        275   280   285
Leu Val Gly Glu Ile Trp Asp Gln Pro Glu Val Val Ala Pro Tyr Tyr
    290   295   300
Gln Ser Leu Asp Ser Thr Phe Asn Phe Asp Leu Ala Tyr Lys Ile Val
305   310   315   320
Asn Ser Val Lys Asn Gly Thr Asp Gln Gly Val Ala Ala Ala Ala Val
                325   330   335
Ala Thr Asp Glu Leu Tyr Lys Thr Tyr Asn Pro Asn Lys Ile Asp Gly
            340   345   350

```
Thr Phe Leu Thr Asn His Asp Gln Asn Arg Val Met Ser Glu Leu Asn
        355                 360                 365
Gly Asp Val Asn Lys Ala Lys Ser Ala Ala Ser Ile Leu Leu Thr Leu
    370                 375                 380
Pro Gly Asn Pro Phe Ile Tyr Tyr Gly Glu Glu Ile Gly Met Thr Gly
385                 390                 395                 400
Gln Lys Pro Asp Glu Leu Ile Arg Glu Pro Phe Arg Trp Tyr Glu Asp
                405                 410                 415
Asp Lys Glu Gly Gln Thr Ser Trp Glu Thr Pro Val Tyr Asn Ile Asp
            420                 425                 430
His Asn Gly Val Ser Val Glu Ala Gln Asp Lys Gln Lys Ala Ser Leu
        435                 440                 445
Leu Ser His Tyr Arg Lys Met Ile Arg Val Arg Gln Gln His Asp Glu
    450                 455                 460
Leu Val Lys Gly Asn Leu Glu Pro Ile Ser Val Asn Asn Ser Gln Val
465                 470                 475                 480
Val Ala Tyr Asn Arg Thr Tyr Lys Asn Lys Ser Ile Gln Val Tyr His
                485                 490                 495
Asn Ile Ser Asp Lys Pro Val Thr Leu Thr Val Ser Asn Lys Gly Lys
            500                 505                 510
Leu Ile Phe Ser Ser Glu Leu Gly Ala Lys Lys Glu Lys Ser Thr Leu
        515                 520                 525
Val Ile Pro Ala Asn Thr Thr Val Leu Val Lys
    530                 535
```

<210> 155
<211> 1773
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 155

```
gtgtcaagaa tgtttgcaaa acgattcaaa acctctttac tgccgttatt cgctggattt      60
ttattgctgt ttcatttggt tctggcagga ccaacggctg cgaatgctga aacggctaac     120
aaatcaaatg agcttacagc accgtcgatc aaaagcggaa ccattcttca tgcttggaat     180
tggtcgttca atacgttaaa acacaatatg aaggatattc atgatgcagg atatacagcg     240
attcagacgt ctccgattaa ccaagtcaag gaagggaacc aaggaaataa aaacatgtcg     300
aactggtact ggctctatca gccgacatcg taccaaattg gcaaccgtta cttaggtact     360
gaacaagaat ttaaagaaat gtgtgcagcc gctgaagaat atggcataaa ggttattgtt     420
gacgcggtca tcaatcatac caccagtgac tatgccgcga tttccaatga gattaagagt     480
attccaaact ggacacatgg aaacacacaa attaaaaact ggtctgatcg atgggatgtc     540
acgcagaatg cattgctcgg gctgtatgac tggaatacac aaaatacaca agtacagtcc     600
tatttgaaac ggttcttaga aagagcattg aatgacgggg cagacggttt tcgatttgat     660
gccgccaaac atatagagct tccggatgat ggcagttacg gcagtcaatt ttggccgaat     720
atcacaaata catctgcaga gttccaatac ggagaaatcc tgcaggatag tgcttcaaga     780
gatgcttcat atgcgaatta tatgaatgtg acagcgtcta actatgggca ttccataagg     840
tccgctttaa agaatcgtaa tctgggcgtg tcgaatatct cccactatgc atcagatgtg     900
tctgcggaca agctagtgac atgggtagaa tcgcatgata cgtatgccaa tgatgatgaa     960
gagtcgacat ggatgagcga tgatgatatc cgtttaggct gggcggtgat agcttctcgt    1020
tcaggcagta cgcctctttt cttttccaga cctgagggag gcggaaatgg tgtgagattc    1080
ccggggaaaa gccaaatagg cgatcgcggg agtgctttat ttgaagatca ggctatcact    1140
gcggtcaata gatttcacaa tgtgatggct ggacagcctg aggaactctc gaacccaaat    1200
ggaaacaacc agatatttat gaatcagcgc ggctcacatg gcgttgtgct ggcaaatgca    1260
ggttcatcct ctgtttctat caatacgcca acaaaattgc ctgatggcag gtatgataat    1320
aaagctgggg caggttcatt tcaagtaaat gacggtaaac tgacaggcac gatcaatgcc    1380
aggtctgtgg ctgtgcttta tcctgatgat attgcaaaag cgcctcatgt tttccttgag    1440
aattacaaaa caggtgtaac acattctttc aatgatcaac tgacgattac actgcgtgca    1500
gatgcgaata caacaaaagc cgtttatcaa atcaataatg gaccagagac ggcgtttaag    1560
gatggagatc aattcacaat cggaaaagga gatccatttg gcaaaacata caccatcatg    1620
ttaaaaggaa cgaacagtga tggtgtaacg aggaccgagg aatacagttt tgttaaaaga    1680
gatccagctt cggccaaaac catcggctat caaaatccga atcattggag ccaggtaaat    1740
gcttatatct ataaacatga tgggggccgg gca                                 1773
```

<210> 156
<211> 591
<212> PRT

<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 156

```
Val Ser Arg Met Phe Ala Lys Arg Phe Lys Thr Ser Leu Leu Pro Leu
 1               5                  10                  15
Phe Ala Gly Phe Leu Leu Leu Phe His Leu Val Leu Ala Gly Pro Thr
            20                  25                  30
Ala Ala Asn Ala Glu Thr Ala Asn Lys Ser Asn Glu Leu Thr Ala Pro
        35                  40                  45
Ser Ile Lys Ser Gly Thr Ile Leu His Ala Trp Asn Trp Ser Phe Asn
    50                  55                  60
Thr Leu Lys His Asn Met Lys Asp Ile His Asp Ala Gly Tyr Thr Ala
65                  70                  75                  80
Ile Gln Thr Ser Pro Ile Asn Gln Val Lys Glu Gly Asn Gln Gly Asn
                85                  90                  95
Lys Asn Met Ser Asn Trp Tyr Trp Leu Tyr Gln Pro Thr Ser Tyr Gln
            100                 105                 110
Ile Gly Asn Arg Tyr Leu Gly Thr Glu Gln Glu Phe Lys Glu Met Cys
        115                 120                 125
Ala Ala Ala Glu Glu Tyr Gly Ile Lys Val Ile Val Asp Ala Val Ile
    130                 135                 140
Asn His Thr Thr Ser Asp Tyr Ala Ala Ile Ser Asn Glu Ile Lys Ser
145                 150                 155                 160
Ile Pro Asn Trp Thr His Gly Asn Thr Gln Ile Lys Asn Trp Ser Asp
                165                 170                 175
Arg Trp Asp Val Thr Gln Asn Ala Leu Leu Gly Leu Tyr Asp Trp Asn
            180                 185                 190
Thr Gln Asn Thr Gln Val Gln Ser Tyr Leu Lys Arg Phe Leu Glu Arg
        195                 200                 205
Ala Leu Asn Asp Gly Ala Asp Gly Phe Arg Phe Asp Ala Ala Lys His
    210                 215                 220
Ile Glu Leu Pro Asp Asp Gly Ser Tyr Gly Ser Gln Phe Trp Pro Asn
225                 230                 235                 240
Ile Thr Asn Thr Ser Ala Glu Phe Gln Tyr Gly Glu Ile Leu Gln Asp
                245                 250                 255
Ser Ala Ser Arg Asp Ala Ser Tyr Ala Asn Tyr Met Asn Val Thr Ala
            260                 265                 270
Ser Asn Tyr Gly His Ser Ile Arg Ser Ala Leu Lys Asn Arg Asn Leu
        275                 280                 285
Gly Val Ser Asn Ile Ser His Tyr Ala Ser Asp Val Ser Ala Asp Lys
    290                 295                 300
Leu Val Thr Trp Val Glu Ser His Asp Thr Tyr Ala Asn Asp Asp Glu
305                 310                 315                 320
Glu Ser Thr Trp Met Ser Asp Asp Asp Ile Arg Leu Gly Trp Ala Val
                325                 330                 335
Ile Ala Ser Arg Ser Gly Ser Thr Pro Leu Phe Phe Ser Arg Pro Glu
            340                 345                 350
Gly Gly Gly Asn Gly Val Arg Phe Pro Gly Lys Ser Gln Ile Gly Asp
        355                 360                 365
Arg Gly Ser Ala Leu Phe Glu Asp Gln Ala Ile Thr Ala Val Asn Arg
    370                 375                 380
Phe His Asn Val Met Ala Gly Gln Pro Glu Glu Leu Ser Asn Pro Asn
385                 390                 395                 400
Gly Asn Asn Gln Ile Phe Met Asn Gln Arg Gly Ser His Gly Val Val
                405                 410                 415
Leu Ala Asn Ala Gly Ser Ser Ser Val Ser Ile Asn Thr Pro Thr Lys
            420                 425                 430
Leu Pro Asp Gly Arg Tyr Asp Asn Lys Ala Gly Ala Gly Ser Phe Gln
        435                 440                 445
Val Asn Asp Gly Lys Leu Thr Gly Thr Ile Asn Ala Arg Ser Val Ala
    450                 455                 460
Val Leu Tyr Pro Asp Asp Ile Ala Lys Ala Pro His Val Phe Leu Glu
465                 470                 475                 480
Asn Tyr Lys Thr Gly Val Thr His Ser Phe Asn Asp Gln Leu Thr Ile
                485                 490                 495
```

Thr Leu Arg Ala Asp Ala Asn Thr Thr Lys Ala Val Tyr Gln Ile Asn
500 505 510
Asn Gly Pro Glu Thr Ala Phe Lys Asp Gly Asp Gln Phe Thr Ile Gly
515 520 525
Lys Gly Asp Pro Phe Gly Lys Thr Tyr Thr Ile Met Leu Lys Gly Thr
530 535 540
Asn Ser Asp Gly Val Thr Arg Thr Glu Glu Tyr Ser Phe Val Lys Arg
545 550 555 560
Asp Pro Ala Ser Ala Lys Thr Ile Gly Tyr Gln Asn Pro Asn His Trp
565 570 575
Ser Gln Val Asn Ala Tyr Ile Tyr Lys His Asp Gly Gly Arg Ala
580 585 590

<210> 157
<211> 1596
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 157
atgcaaacga ttgcaaaaaa aggggatgaa acgatgaaag ggaaaaaatg gacagcatta 60
gctctaacac tgccgctggc tgctagctta tcaacaggcg ttcacgccga aaccgtacat 120
aaaggtaaag ctccaacagc agataaaaac ggtgtctttt atgaggtgta tgtaaactct 180
ttttacgatg caaataaaga tggacatggt gatttaaaag gtcttacaca aaagctggat 240
tatttgaatg acggcaattc tcataccaaa aatgatcttc aagtaaacgg aatttggatg 300
atgccggtaa acccttctcc tagctatcat aaatatgatg taacggacta ttataacatt 360
gatccgcagt acggaaatct gcaagatttt cgcaagctga tgaaagaagc agataaacga 420
gacgtaaagg ttattatgga cctcgttgtg aatcatacaa gcagtgaaca tccttggttt 480
caagctgcat taaaagataa aaacagcaag tacagagatt actatatttg ggccgataaa 540
aatactgatt taaatgaaaa aggatcttgg gggcagcaag tatggcataa agctccaaac 600
ggagagtatt tttatggtac gttttgggaa ggaatgcctg acttaaatta cgataatccc 660
gaagtaagaa aagaaatgat taacgtcggg aaattttggc taaagcaagg cgttgacggg 720
ttccgcttag atgctgcgct tcatattttt aaaggtcaaa cacctgaagg cgctaagaaa 780
aatatcgtgt ggtggaatga gtttagagat gcaatgaaaa aagaaaaccc taacgtatat 840
ctaacgggtg aagtatggga tcaaccggaa gtagtagctc cttactatca atcgcttgat 900
tctttattta actttgattt agcaggaaag attgtaaact ctgtaaaatc aggaaatgat 960
caaggaatcg cgactgcagc agccgcaact gatgagctgt tcaaatcata caatccaaat 1020
aaaattgacg gcattttctt aaccaaccat gaccaaaatc gcgtcatgag tgagctaagc 1080
ggcgatgtga ataaagcaaa gtcagctgcc tctatcttac ttacgcttcc tggcaacccg 1140
tatatttatt acggtgaaga aattggaatg accggtgaaa agcctgatga gttaatccgt 1200
gaaccgttcc gctggtacga aggcaatgga cttggacaaa ccagctggga aacatccgta 1260
tacaacaaag gcggcaatgg tgtgtcagta gagacacaaa caaaacaaaa ggattctttg 1320
ttaaatcatt accgtgaaat gattcgcgtg cgtcagcagc atgaagagtt agtaaaagga 1380
acccttcaat ctatttcagt agacagtaaa gaagtcgttg cctatagccg cacgtataaa 1440
ggcaaatcga ttagcgtgta tcataatatt tcaaatcaac cggtaaaagt atctgtaaca 1500
gcgaaaggta aattgatttt tgctagtgaa aaaggtgcaa aaaaagtcaa aaatcagctt 1560
gtggttccag ctaatacaac ggttttaata aaataa 1596

<210> 158
<211> 531
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 158
Met Gln Thr Ile Ala Lys Lys Gly Asp Glu Thr Met Lys Gly Lys Lys
1 5 10 15
Trp Thr Ala Leu Ala Leu Thr Leu Pro Leu Ala Ala Ser Leu Ser Thr
20 25 30
Gly Val His Ala Glu Thr Val His Lys Gly Lys Ala Pro Thr Ala Asp
35 40 45
Lys Asn Gly Val Phe Tyr Glu Val Tyr Val Asn Ser Phe Tyr Asp Ala
50 55 60
Asn Lys Asp Gly His Gly Asp Leu Lys Gly Leu Thr Gln Lys Leu Asp

65 70 75 80
Tyr Leu Asn Asp Gly Asn Ser His Thr Lys Asn Asp Leu Gln Val Asn
85 90 95
Gly Ile Trp Met Met Pro Val Asn Pro Ser Pro Ser Tyr His Lys Tyr
100 105 110
Asp Val Thr Asp Tyr Tyr Asn Ile Asp Pro Gln Tyr Gly Asn Leu Gln
115 120 125
Asp Phe Arg Lys Leu Met Lys Glu Ala Asp Lys Arg Asp Val Lys Val
130 135 140
Ile Met Asp Leu Val Val Asn His Thr Ser Ser Glu His Pro Trp Phe
145 150 155 160
Gln Ala Ala Leu Lys Asp Lys Asn Ser Lys Tyr Arg Asp Tyr Tyr Ile
165 170 175
Trp Ala Asp Lys Asn Thr Asp Leu Asn Glu Lys Gly Ser Trp Gly Gln
180 185 190
Gln Val Trp His Lys Ala Pro Asn Gly Glu Tyr Phe Tyr Gly Thr Phe
195 200 205
Trp Glu Gly Met Pro Asp Leu Asn Tyr Asp Asn Pro Glu Val Arg Lys
210 215 220
Glu Met Ile Asn Val Gly Lys Phe Trp Leu Lys Gln Gly Val Asp Gly
225 230 235 240
Phe Arg Leu Asp Ala Ala Leu His Ile Phe Lys Gly Gln Thr Pro Glu
245 250 255
Gly Ala Lys Lys Asn Ile Val Trp Trp Asn Glu Phe Arg Asp Ala Met
260 265 270
Lys Lys Glu Asn Pro Asn Val Tyr Leu Thr Gly Glu Val Trp Asp Gln
275 280 285
Pro Glu Val Val Ala Pro Tyr Tyr Gln Ser Leu Asp Ser Leu Phe Asn
290 295 300
Phe Asp Leu Ala Gly Lys Ile Val Asn Ser Val Lys Ser Gly Asn Asp
305 310 315 320
Gln Gly Ile Ala Thr Ala Ala Ala Ala Thr Asp Glu Leu Phe Lys Ser
325 330 335
Tyr Asn Pro Asn Lys Ile Asp Gly Ile Phe Leu Thr Asn His Asp Gln
340 345 350
Asn Arg Val Met Ser Glu Leu Ser Gly Asp Val Asn Lys Ala Lys Ser
355 360 365
Ala Ala Ser Ile Leu Leu Thr Leu Pro Gly Asn Pro Tyr Ile Tyr Tyr
370 375 380
Gly Glu Glu Ile Gly Met Thr Gly Glu Lys Pro Asp Glu Leu Ile Arg
385 390 395 400
Glu Pro Phe Arg Trp Tyr Glu Gly Asn Gly Leu Gly Gln Thr Ser Trp
405 410 415
Glu Thr Ser Val Tyr Asn Lys Gly Gly Asn Gly Val Ser Val Glu Thr
420 425 430
Gln Thr Lys Gln Lys Asp Ser Leu Leu Asn His Tyr Arg Glu Met Ile
435 440 445
Arg Val Arg Gln Gln His Glu Glu Leu Val Lys Gly Thr Leu Gln Ser
450 455 460
Ile Ser Val Asp Ser Lys Glu Val Val Ala Tyr Ser Arg Thr Tyr Lys
465 470 475 480
Gly Lys Ser Ile Ser Val Tyr His Asn Ile Ser Asn Gln Pro Val Lys
485 490 495
Val Ser Val Thr Ala Lys Gly Lys Leu Ile Phe Ala Ser Glu Lys Gly
500 505 510
Ala Lys Lys Val Lys Asn Gln Leu Val Val Pro Ala Asn Thr Thr Val
515 520 525
Leu Ile Lys
530

<210> 159
<211> 1587
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 159
atgcaaaaaa aaggggatga aacgatgaaa gggaaaaaat ggacagcttt agctctaaca 60
ctgccgctgg ctgctagctt atcaacaggc gttcacgccg aaaccgtaca taaaggtaaa 120
tctccaacag cagataaaaa cggtgtattt tatgaggtgt atgtaaactc tttttacgat 180
gcaaataaag atggacatgg tgatttaaaa ggtcttacac aaaagttgga ttatttaaat 240
gatggcaatt ctcatacaaa gaatgatctt caagtaaacg ggatttggat gatgccggtc 300
aacccttctc ccagctatca taaatatgat gtaacggact attataatat tgatccgcag 360
tatggaaatc tgcaagattt tcgcaaactg atgaaagaag cagataaacg agatgtaaaa 420
gtcattatgg acctcgttgt gaatcatacg agcagtgaac acccttggtt tcaagctgca 480
ttaaaagata aaaacagcaa gtacagagat tactatatct gggctgataa aaataccgac 540
ttgaatgaaa aaggatcttg gggacagcaa gtatggcata aagctccaaa cggagagtat 600
ttttacggaa cgttttggga aggaatgccg gacttaaatt acgataatcc tgaagtaaga 660
aaagaaatga ttaacgtagg aaagttttgg ctaaagcaag gagttgatgg gttccgtcta 720
gatgctgcgc ttcatatttt taaaggccaa acacctgaag gcgctaagaa aaatctcctg 780
tggtggaatg aatttagaga tgcaatgaaa aaggaaaacc ctaacgtata tctaacgggt 840
gaagtatggg atcaaccgga agtagtagct ccttactatc aatcgcttga ttctttattt 900
aactttgatt tagcaggaaa gattgtaaac tctgtaaaat caggaaatga tcaaggaatc 960
gcgactgcag cagcggcaac ggatgaactg ttcaaatcat acaatccaaa taaaattgac 1020
ggtattttct taaccaacca tgaccaaaat cgcgtcatga gtgagctaaa cggcgatgtg 1080
aataaagcaa agtcagctgc ctctatctta cttacgcttc ctggcaaccc gtatatttat 1140
tacggtgaag aaatcggcat gaccggtgaa aagcctgatg agttaatccg tgaaccgttc 1200
ccctggtacg aaggaaacgg acttggacaa accagctggg aaacacctgt atataacaaa 1260
ggcggcaacg gcgtgtctgt agaagcacaa acaaaacaaa aggactcttt gttaaatcat 1320
taccgtgaaa tgattcgcgt gcgtcagcag cacgaagagt tagtaaaagg aacgcttcaa 1380
tctatttcag tagacagtaa agaagtcgtt gcctatagcc gtacgtataa aggcaaatcg 1440
attagcgtgt atcataatat ttcaaatcaa ccggtaaaag tatctgtagc agcaaaaggt 1500
aaattgattt ttgctagtga aaaaggtgct aagaaagtca aaaatcagct tgtgattccg 1560
gcgaatacaa cggttttaat aaaataa 1587

<210> 160
<211> 528
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 160
Met Gln Lys Lys Gly Asp Glu Thr Met Lys Gly Lys Lys Trp Thr Ala
 1   5   10   15
Leu Ala Leu Thr Leu Pro Leu Ala Ala Ser Leu Ser Thr Gly Val His
            20   25   30
Ala Glu Thr Val His Lys Gly Lys Ser Pro Thr Ala Asp Lys Asn Gly
        35   40   45
Val Phe Tyr Glu Val Tyr Val Asn Ser Phe Tyr Asp Ala Asn Lys Asp
    50   55   60
Gly His Gly Asp Leu Lys Gly Leu Thr Gln Lys Leu Asp Tyr Leu Asn
65   70   75   80
Asp Gly Asn Ser His Thr Lys Asn Asp Leu Gln Val Asn Gly Ile Trp
                85   90   95
Met Met Pro Val Asn Pro Ser Pro Ser Tyr His Lys Tyr Asp Val Thr
            100   105   110
Asp Tyr Tyr Asn Ile Asp Pro Gln Tyr Gly Asn Leu Gln Asp Phe Arg
        115   120   125
Lys Leu Met Lys Glu Ala Asp Lys Arg Asp Val Lys Val Ile Met Asp
    130   135   140
Leu Val Val Asn His Thr Ser Ser Glu His Pro Trp Phe Gln Ala Ala
145   150   155   160
Leu Lys Asp Lys Asn Ser Lys Tyr Arg Asp Tyr Tyr Ile Trp Ala Asp
                165   170   175
Lys Asn Thr Asp Leu Asn Glu Lys Gly Ser Trp Gly Gln Gln Val Trp
            180   185   190
His Lys Ala Pro Asn Gly Glu Tyr Phe Tyr Gly Thr Phe Trp Glu Gly
        195   200   205
Met Pro Asp Leu Asn Tyr Asp Asn Pro Glu Val Arg Lys Glu Met Ile
    210   215   220
Asn Val Gly Lys Phe Trp Leu Lys Gln Gly Val Asp Gly Phe Arg Leu
225   230   235   240

```
Asp Ala Ala Leu His Ile Phe Lys Gly Gln Thr Pro Glu Gly Ala Lys
                245                 250                 255
Lys Asn Leu Leu Trp Trp Asn Glu Phe Arg Asp Ala Met Lys Lys Glu
            260                 265                 270
Asn Pro Asn Val Tyr Leu Thr Gly Glu Val Trp Asp Gln Pro Glu Val
        275                 280                 285
Val Ala Pro Tyr Tyr Gln Ser Leu Asp Ser Leu Phe Asn Phe Asp Leu
    290                 295                 300
Ala Gly Lys Ile Val Asn Ser Val Lys Ser Gly Asn Asp Gln Gly Ile
305                 310                 315                 320
Ala Thr Ala Ala Ala Ala Thr Asp Glu Leu Phe Lys Ser Tyr Asn Pro
                325                 330                 335
Asn Lys Ile Asp Gly Ile Phe Leu Thr Asn His Asp Gln Asn Arg Val
            340                 345                 350
Met Ser Glu Leu Asn Gly Asp Val Asn Lys Ala Lys Ser Ala Ala Ser
        355                 360                 365
Ile Leu Leu Thr Leu Pro Gly Asn Pro Tyr Ile Tyr Tyr Gly Glu Glu
    370                 375                 380
Ile Gly Met Thr Gly Glu Lys Pro Asp Glu Leu Ile Arg Glu Pro Phe
385                 390                 395                 400
Pro Trp Tyr Glu Gly Asn Gly Leu Gly Gln Thr Ser Trp Glu Thr Pro
                405                 410                 415
Val Tyr Asn Lys Gly Gly Asn Gly Val Ser Val Glu Ala Gln Thr Lys
            420                 425                 430
Gln Lys Asp Ser Leu Leu Asn His Tyr Arg Glu Met Ile Arg Val Arg
        435                 440                 445
Gln Gln His Glu Glu Leu Val Lys Gly Thr Leu Gln Ser Ile Ser Val
    450                 455                 460
Asp Ser Lys Glu Val Val Ala Tyr Ser Arg Thr Tyr Lys Gly Lys Ser
465                 470                 475                 480
Ile Ser Val Tyr His Asn Ile Ser Asn Gln Pro Val Lys Val Ser Val
                485                 490                 495
Ala Ala Lys Gly Lys Leu Ile Phe Ala Ser Glu Lys Gly Ala Lys Lys
            500                 505                 510
Val Lys Asn Gln Leu Val Ile Pro Ala Asn Thr Thr Val Leu Ile Lys
        515                 520                 525
```

<210> 161
<211> 1623
<212> DNA
<213> Bacterial

<400> 161

```
gtggatccaa agaattgtag tcaatttatg caaacgattg caaaaaaagg ggatgaaacg      60
atgaaaggga aaaaatggac agctttagct ctaacactgc cgctggctgc tagcttatca     120
acaggtgttc acgccgaaac cgtacataaa ggtaaagctc caacagcaga taaaaacggt     180
gtcttttatg aggtatatgt aaactctttt tacgatgcaa ataaagatgg acatggtgat     240
ttaaaaggcc ttacacaaaa gttggactat ttaaatgacg gaaattctca tacaaagaat     300
gatcttcaag taaacgggat ttggatgatg ccggtcaacc cttctcctag ctatcataaa     360
tatgatgtaa cggactatta taatattgat ccgcagtatg gaaatctgca agattttcgc     420
aaacttatga aagaagcaga taaacgagac gtaaaagtca ttatggacct tgttgtgaat     480
catacgagca gtgaacaccc ttggtttcaa gctgcgttga aagataaaaa cagcaagtac     540
agagattact atatttgggc tgataaaaat actgacttga atgaaaaagg atcttgggga     600
caacaagtat ggcataaagc tccaaacgga gagtattttt acggaacgtt ctgggaagga     660
atgcctgact taaattacga taaccctgaa gtaagaaaag aaatgattaa cgtcggaaag     720
ttttggctaa aacaaggcgt tgacggcttc cgcttagatg ctgcccttca tatttttaaa     780
ggtcaaacgc ctgaaggcgc taagaaaaac attctatggt ggaatgagtt tagagatgcg     840
atgaaaaaag aaaacccgaa cgtatatcta acgggtgaag tgtgggacca gccagaagta     900
gtagcccctt actatcaatc acttgattct ctatttaatt ttgatttagc aggaaaaatt     960
gtcagctctg taaaagcagg aaatgatcaa ggaatcgcca ctgcagcagc ggcaactgat    1020
gagctgttca aatcatacaa tccaaataaa attgacggca ttttcttaac caaccatgac    1080
caaaatcgcg tcatgagtga gttaagcggc gatgtgaata aagcaaaatc agccgcctct    1140
atcttactta cgcttcctgg aaatccgtat atttattacg gtgaagaaat tggcatgaca    1200
ggtgaaaagc ctgatgaatt aatccgtgaa ccgttccgct ggtacgaagg caacggaatt    1260
ggacaaacta gctgggaaac acctgtatat aacaaaggcg gtaacggcgt gtctgtagaa    1320
gcacaaacaa aacaaaagga ttccttgtta aatcattacc gtgaaatgat tcgtgtgcgc    1380
cagcagcacg aagagttagt aaaaggaacg cttcaatcca tttcagtaga cagtaaagaa    1440
gtcgttgcct atagccgcac gtacaaaggc aaatcgatta gcgtgtatca taatatttca    1500
```

aatcaacctg taaaagtatc tgtagcagcg aaaggtaact tgatttttgc tagtgaaaaa 1560
ggtgctaaga aagtcaaaaa tcagcttgtg attccggcga atgcgacggt tttaataaaa 1620
taa 1623

<210> 162
<211> 540
<212> PRT
<213> Bacterial

<400> 162
Val Asp Pro Lys Asn Cys Ser Gln Phe Met Gln Thr Ile Ala Lys Lys
1 5 10 15
Gly Asp Glu Thr Met Lys Gly Lys Lys Trp Thr Ala Leu Ala Leu Thr
20 25 30
Leu Pro Leu Ala Ala Ser Leu Ser Thr Gly Val His Ala Glu Thr Val
35 40 45
His Lys Gly Lys Ala Pro Thr Ala Asp Lys Asn Gly Val Phe Tyr Glu
50 55 60
Val Tyr Val Asn Ser Phe Tyr Asp Ala Asn Lys Asp Gly His Gly Asp
65 70 75 80
Leu Lys Gly Leu Thr Gln Lys Leu Asp Tyr Leu Asn Asp Gly Asn Ser
85 90 95
His Thr Lys Asn Asp Leu Gln Val Asn Gly Ile Trp Met Met Pro Val
100 105 110
Asn Pro Ser Pro Ser Tyr His Lys Tyr Asp Val Thr Asp Tyr Tyr Asn
115 120 125
Ile Asp Pro Gln Tyr Gly Asn Leu Gln Asp Phe Arg Lys Leu Met Lys
130 135 140
Glu Ala Asp Lys Arg Asp Val Lys Val Ile Met Asp Leu Val Val Asn
145 150 155 160
His Thr Ser Ser Glu His Pro Trp Phe Gln Ala Ala Leu Lys Asp Lys
165 170 175
Asn Ser Lys Tyr Arg Asp Tyr Tyr Ile Trp Ala Asp Lys Asn Thr Asp
180 185 190
Leu Asn Glu Lys Gly Ser Trp Gly Gln Gln Val Trp His Lys Ala Pro
195 200 205
Asn Gly Glu Tyr Phe Tyr Gly Thr Phe Trp Glu Gly Met Pro Asp Leu
210 215 220
Asn Tyr Asp Asn Pro Glu Val Arg Lys Glu Met Ile Asn Val Gly Lys
225 230 235 240
Phe Trp Leu Lys Gln Gly Val Asp Gly Phe Arg Leu Asp Ala Ala Leu
245 250 255
His Ile Phe Lys Gly Gln Thr Pro Glu Gly Ala Lys Lys Asn Ile Leu
260 265 270
Trp Trp Asn Glu Phe Arg Asp Ala Met Lys Lys Glu Asn Pro Asn Val
275 280 285
Tyr Leu Thr Gly Glu Val Trp Asp Gln Pro Glu Val Val Ala Pro Tyr
290 295 300
Tyr Gln Ser Leu Asp Ser Leu Phe Asn Phe Asp Leu Ala Gly Lys Ile
305 310 315 320
Val Ser Ser Val Lys Ala Gly Asn Asp Gln Gly Ile Ala Thr Ala Ala
325 330 335
Ala Ala Thr Asp Glu Leu Phe Lys Ser Tyr Asn Pro Asn Lys Ile Asp
340 345 350
Gly Ile Phe Leu Thr Asn His Asp Gln Asn Arg Val Met Ser Glu Leu
355 360 365
Ser Gly Asp Val Asn Lys Ala Lys Ser Ala Ala Ser Ile Leu Leu Thr
370 375 380
Leu Pro Gly Asn Pro Tyr Ile Tyr Tyr Gly Glu Glu Ile Gly Met Thr
385 390 395 400
Gly Glu Lys Pro Asp Glu Leu Ile Arg Glu Pro Phe Arg Trp Tyr Glu
405 410 415
Gly Asn Gly Ile Gly Gln Thr Ser Trp Glu Thr Pro Val Tyr Asn Lys
420 425 430
Gly Gly Asn Gly Val Ser Val Glu Ala Gln Thr Lys Gln Lys Asp Ser
435 440 445
Leu Leu Asn His Tyr Arg Glu Met Ile Arg Val Arg Gln Gln His Glu
450 455 460

```
Glu Leu Val Lys Gly Thr Leu Gln Ser Ile Ser Val Asp Ser Lys Glu
465                 470                 475                 480
Val Val Ala Tyr Ser Arg Thr Tyr Lys Gly Lys Ser Ile Ser Val Tyr
                485                 490                 495
His Asn Ile Ser Asn Gln Pro Val Lys Val Ser Val Ala Ala Lys Gly
            500                 505                 510
Asn Leu Ile Phe Ala Ser Glu Lys Gly Ala Lys Lys Val Lys Asn Gln
        515                 520                 525
Leu Val Ile Pro Ala Asn Ala Thr Val Leu Ile Lys
    530                 535                 540
```

<210> 163
<211> 1422
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 163

```
atggtacgtc ccgaacgacg ggctgcattg gaaccgacta tcgaacgact cgcagcactt      60
gaaagacatt gggtgacgac ggtgtggatt ccgccggcgt acaaaggcac gtcacagaac     120
gatgtcgggt atggggcgta cgatttatac gatctcggcg aattcaacca aaaagggacg     180
acccggacga agtacgggac gaaagcgcag ctccagaccg ccatctcgaa cttgcgcggt     240
aaagggatcg gtgtgtacgg cgacgtcgtc atgaatcaca agggcggggc cgattatacc     300
gaatccgttc aggcgatcga ggtcaatccg tcgaaccgga accaagaaac gtccggtgag     360
tatggcatct cggcctggac tgggttcaac ttcgcggggc gcaacaatac atactcgccg     420
ttcaaatggc gctggtacca ttttgacggt accgattggg atcagtcacg cagcttgagc     480
cgcatctata agttcaagag cacaggcaag gcgtgggaca cggacgtgtc gaacgagaac     540
ggcaactatg attatcttat gtatgccgac gtcgatttcg agcatcccga ggtccgccaa     600
gagatgaaga actggggcaa atggtacgcc gactcgctcg ggctcgacgg tttccggttg     660
gatgcggtca aacatatcag ccactcgtac ttgaaggagt gggtgacgag cgtgcgccag     720
acgaccggga aagagatgtt cacggtcgcc gagtattgga agaacgatct cggtgccatc     780
aacgactatc tgtataagac gggctacacg cactccgtct tcgatgtgcc gctccattat     840
aacttccaag cggccggtaa cggcggcggg tattacgata tgcgcaacat cttgaaaggc     900
accgtcaccg aacagcatcc gtcgctgtcc gtgacgattg tcgataacca cgactcacag     960
ccgggccagt cgctcgagtc gacggtcgcc aactggttca aaccgctcgc ctacgcgacg    1020
atcatgacgc gcggtcaggg ttatccggcc ctcttctatg gagactatta tggcacgaaa    1080
gggacgacga accgcgaaat cccgaacatg tcgggcacgc tccaaccgat tttgaaggca    1140
cgaaaagact tcgcctacgg gacgcagcat gactacctcg atcatcagga cgtcatcggc    1200
tggacacgtg aaggtgtgac cgaccgtgcc aaatcgggtc tcgcgacgat tctatcggac    1260
ggtccgggcg gctcgaagtg gatgtacgtc ggcaaacaga acgccggcga ggtatggaaa    1320
gacatgacga acaacaacgc ccgtctcgtc acgatcaatg ctgacggctg gggtcagttc    1380
ttcgtcaacg gaggctcggt ctcgatttat acgcaacaat aa                       1422
```

<210> 164
<211> 473
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 164

```
Met Val Arg Pro Glu Arg Arg Ala Ala Leu Glu Pro Thr Ile Glu Arg
 1               5                  10                  15
Leu Ala Ala Leu Glu Arg His Trp Val Thr Thr Val Trp Ile Pro Pro
            20                  25                  30
Ala Tyr Lys Gly Thr Ser Gln Asn Asp Val Gly Tyr Gly Ala Tyr Asp
        35                  40                  45
Leu Tyr Asp Leu Gly Glu Phe Asn Gln Lys Gly Thr Thr Arg Thr Lys
    50                  55                  60
Tyr Gly Thr Lys Ala Gln Leu Gln Thr Ala Ile Ser Asn Leu Arg Gly
65                  70                  75                  80
Lys Gly Ile Gly Val Tyr Gly Asp Val Val Met Asn His Lys Gly Gly
                85                  90                  95
Ala Asp Tyr Thr Glu Ser Val Gln Ala Ile Glu Val Asn Pro Ser Asn
            100                 105                 110
```

```
Arg Asn Gln Glu Thr Ser Gly Glu Tyr Gly Ile Ser Ala Trp Thr Gly
        115                 120                 125
Phe Asn Phe Ala Gly Arg Asn Asn Thr Tyr Ser Pro Phe Lys Trp Arg
    130                 135                 140
Trp Tyr His Phe Asp Gly Thr Asp Trp Asp Gln Ser Arg Ser Leu Ser
145                 150                 155                 160
Arg Ile Tyr Lys Phe Lys Ser Thr Gly Lys Ala Trp Asp Thr Asp Val
                165                 170                 175
Ser Asn Glu Asn Gly Asn Tyr Asp Tyr Leu Met Tyr Ala Asp Val Asp
            180                 185                 190
Phe Glu His Pro Glu Val Arg Gln Glu Met Lys Asn Trp Gly Lys Trp
        195                 200                 205
Tyr Ala Asp Ser Leu Gly Leu Asp Gly Phe Arg Leu Asp Ala Val Lys
    210                 215                 220
His Ile Ser His Ser Tyr Leu Lys Glu Trp Val Thr Ser Val Arg Gln
225                 230                 235                 240
Thr Thr Gly Lys Glu Met Phe Thr Val Ala Glu Tyr Trp Lys Asn Asp
                245                 250                 255
Leu Gly Ala Ile Asn Asp Tyr Leu Tyr Lys Thr Gly Tyr Thr His Ser
            260                 265                 270
Val Phe Asp Val Pro Leu His Tyr Asn Phe Gln Ala Ala Gly Asn Gly
        275                 280                 285
Gly Gly Tyr Tyr Asp Met Arg Asn Ile Leu Lys Gly Thr Val Thr Glu
    290                 295                 300
Gln His Pro Ser Leu Ser Val Thr Ile Val Asp Asn His Asp Ser Gln
305                 310                 315                 320
Pro Gly Gln Ser Leu Glu Ser Thr Val Ala Asn Trp Phe Lys Pro Leu
                325                 330                 335
Ala Tyr Ala Thr Ile Met Thr Arg Gly Gln Gly Tyr Pro Ala Leu Phe
            340                 345                 350
Tyr Gly Asp Tyr Tyr Gly Thr Lys Gly Thr Thr Asn Arg Glu Ile Pro
        355                 360                 365
Asn Met Ser Gly Thr Leu Gln Pro Ile Leu Lys Ala Arg Lys Asp Phe
    370                 375                 380
Ala Tyr Gly Thr Gln His Asp Tyr Leu Asp His Gln Asp Val Ile Gly
385                 390                 395                 400
Trp Thr Arg Glu Gly Val Thr Asp Arg Ala Lys Ser Gly Leu Ala Thr
                405                 410                 415
Ile Leu Ser Asp Gly Pro Gly Gly Ser Lys Trp Met Tyr Val Gly Lys
            420                 425                 430
Gln Asn Ala Gly Glu Val Trp Lys Asp Met Thr Asn Asn Asn Ala Arg
        435                 440                 445
Leu Val Thr Ile Asn Ala Asp Gly Trp Gly Gln Phe Phe Val Asn Gly
    450                 455                 460
Gly Ser Val Ser Ile Tyr Thr Gln Gln
465                 470
```

<210> 165
<211> 1437
<212> DNA
<213> Bacterial

<400> 165

```
atgcagtatt tcgagtggta cgtgccaaat gatggggaac attggaatcg tttgcgtaat      60
gatgctgaaa atttagctca taaaggaatt acatctgtat ggataccacc cgtatataaa     120
ggaacttcac aaaatgatgt agggtatgga gtgtatgatg tatatgattt gggagaattc     180
aatcaaaaag gaacgatacg gacaaaatat gggacaaaag cacaattaaa atctgcaatt     240
gaggctttac ataatcaaaa tatcgatgta tacggtgatg ttgttatgaa ccataaaggt     300
ggggcagatt atactgaggt tgtaacagcc gttgaggtag accgtaacaa tcgaaatatt     360
gaaacatcga gtgattatca aatagatgcg tggacgggat ttgattttcc aggacgcagg     420
gactcctatt ctaattttaa atggagatgg tttcattttg atggaacaga ttgggatgag     480
ggaaggaaat taaatagaat ttataaattt aaaggcgtag gtaaagcttg ggactgggaa     540
gtgtctagtg agaatggtaa ctatgattat ttaatgtatg cagatcttga tttcgatcat     600
cctgaagttg caaatgaaat gaaaaactgg ggaacctggt atgcggacga attaaattta     660
gatggctttc gtttagacgc agttaaacat attgaccatg agtatcttcg tgattgggta     720
aatcatgtta gaaagcaaac ggggaaggaa atgtttacag tagctgaata ttggcaaaat     780
gatattcgta ctttaaacaa ttatttaggg aaagtaaatt ataatcaatc tgtgttcgat     840
gcacctcttc attataattt tcattatgct tcaacaggga atggaaatta tgatatgagg     900
```

```
aatattttaa agggtacggt agtagaaagt catcctacac ttgctgttac tcttgttgag     960
aatcatgatt ctcagcctgg acagtcatta gaatctgttg tgagtccttg gtttaagccg    1020
ttggcctatg catttatttt aacgcgtgca gaagggtatc cttctgtttt ttatggagat    1080
tactatggca caaatggaaa tagtagttat gaaattccaa cgttaaagga taaaattgat    1140
ccaattctga cggcacgaaa aaactttgca tatggtacgc aacatgatta tttagaccat    1200
ccagatgtga ttggctggac aagagaaggg gatagtatac atgctaattc tggtttagca    1260
acattaatct ctgatggacc aggaggatca aaatggatga atgttggaaa gaacaacgca    1320
ggggaaatat ggtacgatat tacgggcaat caaacaaata ctgtaacgat taataaagat    1380
ggatgggggc agttccatgt aaatgggggc tctgtttcaa tatatgttca gaagtaa       1437

<210> 166
<211> 478
<212> PRT
<213> Bacterial

<400> 166
Met Gln Tyr Phe Glu Trp Tyr Val Pro Asn Asp Gly Glu His Trp Asn
 1               5                  10                  15
Arg Leu Arg Asn Asp Ala Glu Asn Leu Ala His Lys Gly Ile Thr Ser
            20                  25                  30
Val Trp Ile Pro Pro Val Tyr Lys Gly Thr Ser Gln Asn Asp Val Gly
        35                  40                  45
Tyr Gly Val Tyr Asp Val Tyr Asp Leu Gly Glu Phe Asn Gln Lys Gly
    50                  55                  60
Thr Ile Arg Thr Lys Tyr Gly Thr Lys Ala Gln Leu Lys Ser Ala Ile
65                  70                  75                  80
Glu Ala Leu His Asn Gln Asn Ile Asp Val Tyr Gly Asp Val Val Met
                85                  90                  95
Asn His Lys Gly Gly Ala Asp Tyr Thr Glu Val Val Thr Ala Val Glu
            100                 105                 110
Val Asp Arg Asn Asn Arg Asn Ile Glu Thr Ser Ser Asp Tyr Gln Ile
        115                 120                 125
Asp Ala Trp Thr Gly Phe Asp Phe Pro Gly Arg Arg Asp Ser Tyr Ser
    130                 135                 140
Asn Phe Lys Trp Arg Trp Phe His Phe Asp Gly Thr Asp Trp Asp Glu
145                 150                 155                 160
Gly Arg Lys Leu Asn Arg Ile Tyr Lys Phe Lys Gly Val Gly Lys Ala
                165                 170                 175
Trp Asp Trp Glu Val Ser Ser Glu Asn Gly Asn Tyr Asp Tyr Leu Met
            180                 185                 190
Tyr Ala Asp Leu Asp Phe Asp His Pro Glu Val Ala Asn Glu Met Lys
        195                 200                 205
Asn Trp Gly Thr Trp Tyr Ala Asp Glu Leu Asn Leu Asp Gly Phe Arg
    210                 215                 220
Leu Asp Ala Val Lys His Ile Asp His Glu Tyr Leu Arg Asp Trp Val
225                 230                 235                 240
Asn His Val Arg Lys Gln Thr Gly Lys Glu Met Phe Thr Val Ala Glu
                245                 250                 255
Tyr Trp Gln Asn Asp Ile Arg Thr Leu Asn Asn Tyr Leu Gly Lys Val
            260                 265                 270
Asn Tyr Asn Gln Ser Val Phe Asp Ala Pro Leu His Tyr Asn Phe His
        275                 280                 285
Tyr Ala Ser Thr Gly Asn Gly Asn Tyr Asp Met Arg Asn Ile Leu Lys
    290                 295                 300
Gly Thr Val Val Glu Ser His Pro Thr Leu Ala Val Thr Leu Val Glu
305                 310                 315                 320
Asn His Asp Ser Gln Pro Gly Gln Ser Leu Glu Ser Val Val Ser Pro
                325                 330                 335
Trp Phe Lys Pro Leu Ala Tyr Ala Phe Ile Leu Thr Arg Ala Glu Gly
            340                 345                 350
Tyr Pro Ser Val Phe Tyr Gly Asp Tyr Tyr Gly Thr Asn Gly Asn Ser
        355                 360                 365
Ser Tyr Glu Ile Pro Thr Leu Lys Asp Lys Ile Asp Pro Ile Leu Thr
    370                 375                 380
Ala Arg Lys Asn Phe Ala Tyr Gly Thr Gln His Asp Tyr Leu Asp His
385                 390                 395                 400
Pro Asp Val Ile Gly Trp Thr Arg Glu Gly Asp Ser Ile His Ala Asn
                405                 410                 415
```

```
Ser Gly Leu Ala Thr Leu Ile Ser Asp Gly Pro Gly Gly Ser Lys Trp
            420                 425                 430
Met Asn Val Gly Lys Asn Asn Ala Gly Glu Ile Trp Tyr Asp Ile Thr
        435                 440                 445
Gly Asn Gln Thr Asn Thr Val Thr Ile Asn Lys Asp Gly Trp Gly Gln
    450                 455                 460
Phe His Val Asn Gly Gly Ser Val Ser Ile Tyr Val Gln Lys
465                 470                 475
```

<210> 167
<211> 1596
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 167

```
atgcaaacga ttgcaaaaaa aggggatgaa acgatgaaag ggaaaaaatg gacagcttta      60
gctctaacac tgccgctggc tgctagctta tcaacaggcg ttcacgccga aaccgtacat     120
aaaggtaaat ctccaacagc agataaaaac ggtgtatttt atgaggtgta tgtaaactct     180
ttttacgatg caaataaaga tggacatggt gatttaaaag gtcttacaca aaagttggat     240
tatttaaatg atggcaattc tcatacaaag aatgatcttc aagtaaacgg gatttggatg     300
atgccggtca acccttctcc cagctatcat aaatatgatg taacggacta ttataatatt     360
gatccgcagt atggaaatct gcaagatttt cgcaaactga tgaaagaagc agataaacga     420
gatgtaaaag tcattatgga cctcgttgtg aatcatacga gcagtgaaca cccttggttt     480
caagctgcat taaaagataa aaacagcaag tacagagatt actatatctg ggctgataaa     540
aataccgact tgaatgaaaa aggatcttgg ggacagcaag tatggcataa agccccaaac     600
ggagagtatt tttacggaac gttttgggaa ggaatgccgg acttaaatta cgataatcct     660
gaagtaagaa aagaaatgat taacgtagga aagttttggc taaagcaagg agttgacggg     720
ttccgtctag atgctgcgct tcatattttt aaaggccaaa cacctgaagg cgctaagaaa     780
aatctcctgt ggtggaatga atttagagat gcaatgaaaa aggaaaaccc taacgtatat     840
ctaacgggtg aagtatggga tcaaccggaa gtagtagctc cttactatca atcgcttgat     900
tctttattta actttgattt agcaggaaag attgtaaact ctgtaaaatc aggaaatgat     960
caaggaatcg cgactgcagc agcggcaacg gatgaactgt tcaaatcata caatccaaat    1020
aaaattgacg gtattttctt aaccaaccat gaccaaaatc gcgtcatgag tgagctaagc    1080
ggcgatgtga ataaagcaaa gtcagctgcc tctatcttac ttacgcttcc tggcaacccg    1140
tatatttatt acggtgaaga aatcggcatg accggtgaaa agcctgatga gttaatccgt    1200
gaaccgttcc gctggtacga aggaaacgga cttggacaaa ccagctggga aacacctgta    1260
tacaacaaag gcggcaacgg cgtgtctgta gaagcacaaa caaaacaaaa ggactctttg    1320
ttaaatcatt accgtgaaat gattcgcgtg cgtcagcagc acgaagagtt agtaaaagga    1380
acgcttcaat ctatttcagt agacagtaaa gaagtcgttg cctatagccg cacgtataaa    1440
ggcaaatcga ttagcgtgta tcataatatt tcaaatcaac cggtaaaagt atctgtagca    1500
gcaaaaggta aattgatttt tggtagtgaa aaaggtgcta agaaagtcaa aaatcagctt    1560
gtgattccgg cgaatacaac ggttttaata aaataa                              1596
```

<210> 168
<211> 531
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 168

```
Met Gln Thr Ile Ala Lys Lys Gly Asp Glu Thr Met Lys Gly Lys Lys
 1               5                  10                  15
Trp Thr Ala Leu Ala Leu Thr Leu Pro Leu Ala Ala Ser Leu Ser Thr
            20                  25                  30
Gly Val His Ala Glu Thr Val His Lys Gly Lys Ser Pro Thr Ala Asp
        35                  40                  45
Lys Asn Gly Val Phe Tyr Glu Val Tyr Val Asn Ser Phe Tyr Asp Ala
    50                  55                  60
Asn Lys Asp Gly His Gly Asp Leu Lys Gly Leu Thr Gln Lys Leu Asp
65                  70                  75                  80
Tyr Leu Asn Asp Gly Asn Ser His Thr Lys Asn Asp Leu Gln Val Asn
                85                  90                  95
Gly Ile Trp Met Met Pro Val Asn Pro Ser Pro Ser Tyr His Lys Tyr
```

```
            100                 105                 110
Asp Val Thr Asp Tyr Tyr Asn Ile Asp Pro Gln Tyr Gly Asn Leu Gln
        115                 120                 125
Asp Phe Arg Lys Leu Met Lys Glu Ala Asp Lys Arg Asp Val Lys Val
    130                 135                 140
Ile Met Asp Leu Val Val Asn His Thr Ser Ser Glu His Pro Trp Phe
145                 150                 155                 160
Gln Ala Ala Leu Lys Asp Lys Asn Ser Lys Tyr Arg Asp Tyr Tyr Ile
                165                 170                 175
Trp Ala Asp Lys Asn Thr Asp Leu Asn Glu Lys Gly Ser Trp Gly Gln
            180                 185                 190
Gln Val Trp His Lys Ala Pro Asn Gly Glu Tyr Phe Tyr Gly Thr Phe
        195                 200                 205
Trp Glu Gly Met Pro Asp Leu Asn Tyr Asp Asn Pro Glu Val Arg Lys
    210                 215                 220
Glu Met Ile Asn Val Gly Lys Phe Trp Leu Lys Gln Gly Val Asp Gly
225                 230                 235                 240
Phe Arg Leu Asp Ala Ala Leu His Ile Phe Lys Gly Gln Thr Pro Glu
                245                 250                 255
Gly Ala Lys Lys Asn Leu Leu Trp Trp Asn Glu Phe Arg Asp Ala Met
            260                 265                 270
Lys Lys Glu Asn Pro Asn Val Tyr Leu Thr Gly Glu Val Trp Asp Gln
        275                 280                 285
Pro Glu Val Val Ala Pro Tyr Tyr Gln Ser Leu Asp Ser Leu Phe Asn
    290                 295                 300
Phe Asp Leu Ala Gly Lys Ile Val Asn Ser Val Lys Ser Gly Asn Asp
305                 310                 315                 320
Gln Gly Ile Ala Thr Ala Ala Ala Ala Thr Asp Glu Leu Phe Lys Ser
                325                 330                 335
Tyr Asn Pro Asn Lys Ile Asp Gly Ile Phe Leu Thr Asn His Asp Gln
            340                 345                 350
Asn Arg Val Met Ser Glu Leu Ser Gly Asp Val Asn Lys Ala Lys Ser
        355                 360                 365
Ala Ala Ser Ile Leu Leu Thr Leu Pro Gly Asn Pro Tyr Ile Tyr Tyr
    370                 375                 380
Gly Glu Glu Ile Gly Met Thr Gly Glu Lys Pro Asp Glu Leu Ile Arg
385                 390                 395                 400
Glu Pro Phe Arg Trp Tyr Glu Gly Asn Gly Leu Gly Gln Thr Ser Trp
                405                 410                 415
Glu Thr Pro Val Tyr Asn Lys Gly Gly Asn Gly Val Ser Val Glu Ala
            420                 425                 430
Gln Thr Lys Gln Lys Asp Ser Leu Leu Asn His Tyr Arg Glu Met Ile
        435                 440                 445
Arg Val Arg Gln Gln His Glu Glu Leu Val Lys Gly Thr Leu Gln Ser
    450                 455                 460
Ile Ser Val Asp Ser Lys Glu Val Val Ala Tyr Ser Arg Thr Tyr Lys
465                 470                 475                 480
Gly Lys Ser Ile Ser Val Tyr His Asn Ile Ser Asn Gln Pro Val Lys
                485                 490                 495
Val Ser Val Ala Ala Lys Gly Lys Leu Ile Phe Gly Ser Glu Lys Gly
            500                 505                 510
Ala Lys Lys Val Lys Asn Gln Leu Val Ile Pro Ala Asn Thr Thr Val
        515                 520                 525
Leu Ile Lys
    530

<210> 169
<211> 1524
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 169
atgaaaacat tcaaattaaa acgcactttt ttaccgctaa ccttgctgct cagtgctcct      60
gcctttgctg ggcaaaatgg caccatgatg cagtattttc attggtatgt acctaatgat     120
ggcgcattat ggacgcaggt tgaaagcaat gctccagcac tcgctgaaaa cggttttaca     180
```

```
gcgctctggc taccgccagc ttacaaaggc gcgggcggca gtaatgacgt cggttatggc     240
gtctatgata tgtacgattt aggtgagttt gatcaaaaag gctcagtacg aaccaaatac     300
ggcaccaagg ctcagtacat ctctgcaatc aatgccgcgc acaacaacaa tatccaaatc     360
tacggcgatg ttgtgtttaa ccaccgaggt ggtgctgatg ggaagtcgtg ggtcgatacc     420
aagcgcgttg attgggacaa ccgtaacatt gaactgggcg acaaatggat tgaagcttgg     480
gttgagttta attttcctgg ccgcaacgac aaatactcaa acttccattg gacttggtat     540
cactttgacg gtgttgactg ggatgatgcc ggcaaagaaa aagcgatctt taaattcaaa     600
ggcgaaggaa aagcatggga ttgggaagtc agctctgaaa aaggcaatta cgactaccta     660
atgtacgccg atttagacat ggatcaccaa gaagttaaac aagagctgaa agattggggt     720
gagtggtaca tcaacatgac cggcgttgat ggctttagaa tggatgccgt gaagcacatt     780
aaatatcagt atctacaaga gtggattgat catttacgtt ggaaaacagg caaagagctt     840
ttcaccgttg gtgagtattg gaattacgac gtaaatcaac tgcataactt tattactaag     900
acctctggca gtatgtcgtt gttcgatgcg ccgcttcaca tgaacttcta caacgcgtca     960
aaatctggcg gcaattacga tatgcgccaa atcatgaatg gcacgttgat gaaggacaac    1020
ccagtcaaag ctgtgactct cgtagaaaac cacgatacac agccattgca ggcgttagag    1080
tcgacagtgg attggtggtt caagcctctt gcttacgcat tcattttatt gcgtgaagaa    1140
ggttatccat cagtgttcta cgcagattac tacggcgcgc agtacagcga caaaggctac    1200
aacatcaata tggccaaagt tccttacatt gaagaacttg taacactgcg taaagagtat    1260
gcgtatggca aacagaattc ttatctcgac cactgggatg tgattggctg gacccgagag    1320
ggcgatgctg aacatccaaa ctcaatggcg gtgatcatga gtgatggacc aggtggcaaa    1380
aaatggatgt ataccggtaa gccaagcacg cgctatgtcg acaagctggg tatccgaact    1440
gaagaagttt ggaccgatac caatggctgg gcagaatttc ctgtcaatgg tggttcagtc    1500
tcggtttggg tgggcgttaa gtaa                                           1524

<210> 170
<211> 507
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 170
Met Lys Thr Phe Lys Leu Lys Arg Thr Phe Leu Pro Leu Thr Leu Leu
 1               5                  10                  15
Leu Ser Ala Pro Ala Phe Ala Gly Gln Asn Gly Thr Met Met Gln Tyr
            20                  25                  30
Phe His Trp Tyr Val Pro Asn Asp Gly Ala Leu Trp Thr Gln Val Glu
        35                  40                  45
Ser Asn Ala Pro Ala Leu Ala Glu Asn Gly Phe Thr Ala Leu Trp Leu
    50                  55                  60
Pro Pro Ala Tyr Lys Gly Ala Gly Gly Ser Asn Asp Val Gly Tyr Gly
65                  70                  75                  80
Val Tyr Asp Met Tyr Asp Leu Gly Glu Phe Asp Gln Lys Gly Ser Val
                85                  90                  95
Arg Thr Lys Tyr Gly Thr Lys Ala Gln Tyr Ile Ser Ala Ile Asn Ala
            100                 105                 110
Ala His Asn Asn Asn Ile Gln Ile Tyr Gly Asp Val Val Phe Asn His
        115                 120                 125
Arg Gly Gly Ala Asp Gly Lys Ser Trp Val Asp Thr Lys Arg Val Asp
    130                 135                 140
Trp Asp Asn Arg Asn Ile Glu Leu Gly Asp Lys Trp Ile Glu Ala Trp
145                 150                 155                 160
Val Glu Phe Asn Phe Pro Gly Arg Asn Asp Lys Tyr Ser Asn Phe His
                165                 170                 175
Trp Thr Trp Tyr His Phe Asp Gly Val Asp Trp Asp Asp Ala Gly Lys
            180                 185                 190
Glu Lys Ala Ile Phe Lys Phe Lys Gly Glu Gly Lys Ala Trp Asp Trp
        195                 200                 205
Glu Val Ser Ser Glu Lys Gly Asn Tyr Asp Tyr Leu Met Tyr Ala Asp
    210                 215                 220
Leu Asp Met Asp His Gln Glu Val Lys Gln Glu Leu Lys Asp Trp Gly
225                 230                 235                 240
Glu Trp Tyr Ile Asn Met Thr Gly Val Asp Gly Phe Arg Met Asp Ala
                245                 250                 255
Val Lys His Ile Lys Tyr Gln Tyr Leu Gln Glu Trp Ile Asp His Leu
            260                 265                 270
Arg Trp Lys Thr Gly Lys Glu Leu Phe Thr Val Gly Glu Tyr Trp Asn
```

```
        275                 280                 285
Tyr Asp Val Asn Gln Leu His Asn Phe Ile Thr Lys Thr Ser Gly Ser
    290                 295                 300
Met Ser Leu Phe Asp Ala Pro Leu His Met Asn Phe Tyr Asn Ala Ser
305                 310                 315                 320
Lys Ser Gly Gly Asn Tyr Asp Met Arg Gln Ile Met Asn Gly Thr Leu
                325                 330                 335
Met Lys Asp Asn Pro Val Lys Ala Val Thr Leu Val Glu Asn His Asp
            340                 345                 350
Thr Gln Pro Leu Gln Ala Leu Glu Ser Thr Val Asp Trp Trp Phe Lys
        355                 360                 365
Pro Leu Ala Tyr Ala Phe Ile Leu Leu Arg Glu Glu Gly Tyr Pro Ser
    370                 375                 380
Val Phe Tyr Ala Asp Tyr Tyr Gly Ala Gln Tyr Ser Asp Lys Gly Tyr
385                 390                 395                 400
Asn Ile Asn Met Ala Lys Val Pro Tyr Ile Glu Glu Leu Val Thr Leu
                405                 410                 415
Arg Lys Glu Tyr Ala Tyr Gly Lys Gln Asn Ser Tyr Leu Asp His Trp
            420                 425                 430
Asp Val Ile Gly Trp Thr Arg Glu Gly Asp Ala Glu His Pro Asn Ser
        435                 440                 445
Met Ala Val Ile Met Ser Asp Gly Pro Gly Gly Lys Lys Trp Met Tyr
    450                 455                 460
Thr Gly Lys Pro Ser Thr Arg Tyr Val Asp Lys Leu Gly Ile Arg Thr
465                 470                 475                 480
Glu Glu Val Trp Thr Asp Thr Asn Gly Trp Ala Glu Phe Pro Val Asn
                485                 490                 495
Gly Gly Ser Val Ser Val Trp Val Gly Val Lys
            500                 505
```

<210> 171
<211> 1431
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 171

```
gtgtatgtaa actcttttta cgatgcaaat aaagatggac atggtgattt aaaaggtctt      60
acacaaaagt tggattattt aaatgatggc aattctcata caaagaatga tcttcaagta     120
aacgggattt ggatgatgcc ggtcaaccct tctcccagct atcataaata tgatgtaacg     180
gactattata atattgatcc gcagtatgga aatctgcaag attttcgcaa actgatgaaa     240
gaagcagata aacgagatgt aaaagtcatt atggacctcg ttgtgaatca tacgagcagt     300
gaacaccctt ggtttcaagc tgcattaaaa gataaaaaca gcaagtacag agattactat     360
atctgggctg ataaaaatac cgacttgaat gaaaaaggat cttggggaca gcaagtatgg     420
cataaagccc caaacggaga gtatttttac ggaacgtttt gggaaggaat gccggactta     480
aattacgata atcctgaagt aagaaaagaa atgattaacg taggaaagtt ttggctaaag     540
caaggagttg acgggttccg tctagatgct gcgcttcata tttttaaagg ccaaacacct     600
gaaggcgcta agaaaaatct cctgtggtgg aatgaattta gagatgcaat gaaaaaggaa     660
aaccctaacg tatatctaac gggtgaagta tgggatcaac cggaagtagt agctccttac     720
tatcaatcgc ttgattcttt atttaacttt gatttagcag gaaagattgt aaactctgta     780
aaatcaggaa atgatcaagg aatcgcgact gcagcagcgg caacggatga actgttcaaa     840
tcatacaatc caaataaaat tgacggtatt ttcttaacca accatgacca aaatcgcgtc     900
atgagtgagc taagcggcga tgtgaataaa gcaaagtcag ctgcctctat cttacttacg     960
cttcctggca acccgtatat ttattacggt gaagaaatcg gcatgaccgg tgaaaagcct    1020
gatgagttaa tccgtgaacc gttccgctgg tacgaaggaa acggacttgg acaaaccagc    1080
tgggaaacac ctgtatacaa caaaggcggc aacggcgtgt ctgtagaagc acaaacaaaa    1140
caaaaggact ctttgttaaa tcattaccgt gaaatgattc gcgtgcgtca gcagcacgaa    1200
gagttagtaa aaggaacgct tcaatctatt tcagtagaca gtaaagaagt cgttgcctat    1260
agccgcacgt ataaaggcaa atcgattagc gtgtatcata atatttcaaa tcaaccggta    1320
aaagtatctg tagcagcaaa aggtaaattg atttttggta gtgaaaaagg tgctaagaaa    1380
gtcaaaaatc agcttgtgat tccggcgaat acaacggttt taataaaata a             1431
```

<210> 172
<211> 476
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 172
Val Tyr Val Asn Ser Phe Tyr Asp Ala Asn Lys Asp Gly His Gly Asp
 1   5   10   15
Leu Lys Gly Leu Thr Gln Lys Leu Asp Tyr Leu Asn Asp Gly Asn Ser
            20   25   30
His Thr Lys Asn Asp Leu Gln Val Asn Gly Ile Trp Met Met Pro Val
        35   40   45
Asn Pro Ser Pro Ser Tyr His Lys Tyr Asp Val Thr Asp Tyr Tyr Asn
    50   55   60
Ile Asp Pro Gln Tyr Gly Asn Leu Gln Asp Phe Arg Lys Leu Met Lys
65   70   75   80
Glu Ala Asp Lys Arg Asp Val Lys Val Ile Met Asp Leu Val Val Asn
                85   90   95
His Thr Ser Ser Glu His Pro Trp Phe Gln Ala Ala Leu Lys Asp Lys
            100   105   110
Asn Ser Lys Tyr Arg Asp Tyr Tyr Ile Trp Ala Asp Lys Asn Thr Asp
        115   120   125
Leu Asn Glu Lys Gly Ser Trp Gly Gln Gln Val Trp His Lys Ala Pro
    130   135   140
Asn Gly Glu Tyr Phe Tyr Gly Thr Phe Trp Glu Gly Met Pro Asp Leu
145   150   155   160
Asn Tyr Asp Asn Pro Glu Val Arg Lys Glu Met Ile Asn Val Gly Lys
                165   170   175
Phe Trp Leu Lys Gln Gly Val Asp Gly Phe Arg Leu Asp Ala Ala Leu
            180   185   190
His Ile Phe Lys Gly Gln Thr Pro Glu Gly Ala Lys Lys Asn Leu Leu
        195   200   205
Trp Trp Asn Glu Phe Arg Asp Ala Met Lys Lys Glu Asn Pro Asn Val
    210   215   220
Tyr Leu Thr Gly Glu Val Trp Asp Gln Pro Glu Val Val Ala Pro Tyr
225   230   235   240
Tyr Gln Ser Leu Asp Ser Leu Phe Asn Phe Asp Leu Ala Gly Lys Ile
                245   250   255
Val Asn Ser Val Lys Ser Gly Asn Asp Gln Gly Ile Ala Thr Ala Ala
            260   265   270
Ala Ala Thr Asp Glu Leu Phe Lys Ser Tyr Asn Pro Asn Lys Ile Asp
        275   280   285
Gly Ile Phe Leu Thr Asn His Asp Gln Asn Arg Val Met Ser Glu Leu
    290   295   300
Ser Gly Asp Val Asn Lys Ala Lys Ser Ala Ala Ser Ile Leu Leu Thr
305   310   315   320
Leu Pro Gly Asn Pro Tyr Ile Tyr Tyr Gly Glu Glu Ile Gly Met Thr
                325   330   335
Gly Glu Lys Pro Asp Glu Leu Ile Arg Glu Pro Phe Arg Trp Tyr Glu
            340   345   350
Gly Asn Gly Leu Gly Gln Thr Ser Trp Glu Thr Pro Val Tyr Asn Lys
        355   360   365
Gly Gly Asn Gly Val Ser Val Glu Ala Gln Thr Lys Gln Lys Asp Ser
    370   375   380
Leu Leu Asn His Tyr Arg Glu Met Ile Arg Val Arg Gln Gln His Glu
385   390   395   400
Glu Leu Val Lys Gly Thr Leu Gln Ser Ile Ser Val Asp Ser Lys Glu
                405   410   415
Val Val Ala Tyr Ser Arg Thr Tyr Lys Gly Lys Ser Ile Ser Val Tyr
            420   425   430
His Asn Ile Ser Asn Gln Pro Val Lys Val Ser Val Ala Ala Lys Gly
        435   440   445
Lys Leu Ile Phe Gly Ser Glu Lys Gly Ala Lys Lys Val Lys Asn Gln
    450   455   460
Leu Val Ile Pro Ala Asn Thr Thr Val Leu Ile Lys
465   470   475

<210> 173
<211> 1596

<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 173
atgcaaacga ttgcaaaaaa aggggatgaa acgatgaaag ggaaaaaatg gacagcttta 60
gctctaacac tgccgctggc tgctagctta tcaacaggcg ttcacgcaga aactgtacat 120
aaaggtaaag ctccaacagc agataaaaac ggtgtttttt atgaggtgta tgtaaactct 180
ttttacgatg caaataaaga tggacatggt gatttaaaag gtctgacaca aaagttggat 240
tatttaaatg acggcaattc tcatacaaag aatgatcttc aagtaaacgg gatttggatg 300
atgccggtaa acccttctcc tagctatcat aaatatgatg taacggacta ttataacatt 360
gatcctcagt acggaagtct gcaagatttc cgcaaactga tgaaagaagc agataaacga 420
gacgtaaaag ttattatgga ccttgttgtg aatcatacga gcagtgaaca cccttggttt 480
caagctgcac taaaagataa aaacagcaag tacagagatt actatatttg ggctgataaa 540
aataccgatt tgaatgaaaa aggatcttgg ggacagcaag tatggcataa agctccaaac 600
ggagagtatt tttacggaac gttctgggaa ggaatgcctg acttaaatta cgataaccct 660
gaagtaagaa aagaaatgat taacgtcgga aagttttggc taaagcaagg cgttgatggc 720
ttccgcttag atgctgccct tcatatcttt aaaggtcaaa ctcctgaagg cgctaagaaa 780
aatctcctgt ggtggaatga gtttagagat gcaatgaaaa aagaaaaccc taacgtatat 840
ctaacgggtg aagtatggga tcagccggaa gtagtagctc cttattatca atcgcttgat 900
tccctattta actttgattt agcaggaaaa attgtcagct ctgtaaaagc aggaaatgat 960
caaggaatcg ccactgcagc agcggcaacg gatgagctgt tcaaatcata caatccaaat 1020
aaaattgacg gcattttctt aaccaaccat gaccaaaacc gcgtcatgag tgagctaagc 1080
ggagatgtga ataaagcaaa atcagctgct tctatcttac ttacgcttcc tggaaatccg 1140
tatatttatt acggtgaaga aattggcatg accggtgaaa agcctgatga attaatccgt 1200
gaaccgttcc gctggtacga aggcaacgga attggacaaa ctagctggga aacacctgta 1260
tataacaaag gcggcaatgg tgtgtctgta gaagcacaaa ccaaacaaaa ggattctttg 1320
ttaaatcatt accgtgaaat gattcgcgtg cgtcagcagc acgaagagtt agtaaaagga 1380
acgcttcagt ctatttcagt agacagtaaa gaagttgtcg cttatagccg tacgtataaa 1440
ggcaactcca ttagtgtgta tcataatatt tcaaatcaac ctgtaaaagt atctgtagcg 1500
gcgaaaggta aattgatttt tgctagtgaa aaaggtgcta aaaaaggcaa aaatcagctt 1560
gtgattccgg cgaatgcgac ggttttaata aaataa 1596

<210> 174
<211> 531
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 174
Met Gln Thr Ile Ala Lys Lys Gly Asp Glu Thr Met Lys Gly Lys Lys
 1   5   10   15
Trp Thr Ala Leu Ala Leu Thr Leu Pro Leu Ala Ala Ser Leu Ser Thr
            20   25   30
Gly Val His Ala Glu Thr Val His Lys Gly Lys Ala Pro Thr Ala Asp
        35   40   45
Lys Asn Gly Val Phe Tyr Glu Val Tyr Val Asn Ser Phe Tyr Asp Ala
    50   55   60
Asn Lys Asp Gly His Gly Asp Leu Lys Gly Leu Thr Gln Lys Leu Asp
65   70   75   80
Tyr Leu Asn Asp Gly Asn Ser His Thr Lys Asn Asp Leu Gln Val Asn
                85   90   95
Gly Ile Trp Met Met Pro Val Asn Pro Ser Pro Ser Tyr His Lys Tyr
            100   105   110
Asp Val Thr Asp Tyr Tyr Asn Ile Asp Pro Gln Tyr Gly Ser Leu Gln
        115   120   125
Asp Phe Arg Lys Leu Met Lys Glu Ala Asp Lys Arg Asp Val Lys Val
    130   135   140
Ile Met Asp Leu Val Val Asn His Thr Ser Ser Glu His Pro Trp Phe
145   150   155   160
Gln Ala Ala Leu Lys Asp Lys Asn Ser Lys Tyr Arg Asp Tyr Tyr Ile
                165   170   175
Trp Ala Asp Lys Asn Thr Asp Leu Asn Glu Lys Gly Ser Trp Gly Gln
            180   185   190

```
Gln Val Trp His Lys Ala Pro Asn Gly Glu Tyr Phe Tyr Gly Thr Phe
        195                 200                 205
Trp Glu Gly Met Pro Asp Leu Asn Tyr Asp Asn Pro Glu Val Arg Lys
    210                 215                 220
Glu Met Ile Asn Val Gly Lys Phe Trp Leu Lys Gln Gly Val Asp Gly
225                 230                 235                 240
Phe Arg Leu Asp Ala Ala Leu His Ile Phe Lys Gly Gln Thr Pro Glu
                245                 250                 255
Gly Ala Lys Lys Asn Leu Leu Trp Trp Asn Glu Phe Arg Asp Ala Met
            260                 265                 270
Lys Lys Glu Asn Pro Asn Val Tyr Leu Thr Gly Glu Val Trp Asp Gln
        275                 280                 285
Pro Glu Val Val Ala Pro Tyr Tyr Gln Ser Leu Asp Ser Leu Phe Asn
    290                 295                 300
Phe Asp Leu Ala Gly Lys Ile Val Ser Ser Val Lys Ala Gly Asn Asp
305                 310                 315                 320
Gln Gly Ile Ala Thr Ala Ala Ala Ala Thr Asp Glu Leu Phe Lys Ser
                325                 330                 335
Tyr Asn Pro Asn Lys Ile Asp Gly Ile Phe Leu Thr Asn His Asp Gln
            340                 345                 350
Asn Arg Val Met Ser Glu Leu Ser Gly Asp Val Asn Lys Ala Lys Ser
        355                 360                 365
Ala Ala Ser Ile Leu Leu Thr Leu Pro Gly Asn Pro Tyr Ile Tyr Tyr
    370                 375                 380
Gly Glu Glu Ile Gly Met Thr Gly Glu Lys Pro Asp Glu Leu Ile Arg
385                 390                 395                 400
Glu Pro Phe Arg Trp Tyr Glu Gly Asn Gly Ile Gly Gln Thr Ser Trp
                405                 410                 415
Glu Thr Pro Val Tyr Asn Lys Gly Gly Asn Gly Val Ser Val Glu Ala
            420                 425                 430
Gln Thr Lys Gln Lys Asp Ser Leu Leu Asn His Tyr Arg Glu Met Ile
        435                 440                 445
Arg Val Arg Gln Gln His Glu Glu Leu Val Lys Gly Thr Leu Gln Ser
    450                 455                 460
Ile Ser Val Asp Ser Lys Glu Val Val Ala Tyr Ser Arg Thr Tyr Lys
465                 470                 475                 480
Gly Asn Ser Ile Ser Val Tyr His Asn Ile Ser Asn Gln Pro Val Lys
                485                 490                 495
Val Ser Val Ala Ala Lys Gly Lys Leu Ile Phe Ala Ser Glu Lys Gly
            500                 505                 510
Ala Lys Lys Gly Lys Asn Gln Leu Val Ile Pro Ala Asn Ala Thr Val
        515                 520                 525
Leu Ile Lys
    530

<210> 175
<211> 1398
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 175
atgaaaaata taatacgact ttgtgctgcc agcgctatcc tcacggtgtc ccacgccagt      60
tacgccgacg caattttaca cgcgtttaac tggcaatata ccgatgtaac cgccaatgca     120
aatcaaattg ccgcaaatgg ctttaaaaaa gtcctcattt cacccgcaat gaaatccagc     180
ggcagtcaat ggtgggcccg ctatcaaccg caagacttgc gtgtcattga ttctccgctg     240
ggcaacaaac aagatttagt cgcgatgatc aatgcgctca acagcgttgg ggtcgacgtg     300
tatgctgacg tggtgcttaa ccatatggct aacgagtcat ggaagcgcag tgacctgaac     360
tacccgggga gtgaggtgct caacgactat caatcccgca gtgcttacta tcaaaggcaa     420
acacttttcg gcaatttaca ggagaacctt ttttccgaga atgatttcca tccggcaggc     480
tgtattacca attggaatga tcctggccac gtccagtatt ggcgcttgtg cggcggacag     540
ggcgatactg ggctaccgga tctcgatcct aatcaatggg ttgtgagtca gcagaagagt     600
tacttgaacg cactcaaatc aatgggaatc aaagggttcc gtatcgatgc ggtcaaacat     660
atgagtcaat atcaaataga ccaagtgttt accccagaca ttaccgctgg tatgcatata     720
ttcggagaag tcattaccag tggtgggcaa ggtgatagcg gctatgaggc ttttcttgcc     780
ccttacctta ataataccga tcacgccgct tatgacttcc cgctatttgc atcgattcga     840
```

gccgcgtttt cattctctgg tgggttaaat cagctacaca atccacaagc ctatggccaa 900
gcgttacagg actcacgtgc gatcaccttt acgattaccc acgacattcc aaccaatgac 960
ggtttccgct accagatcat ggatccaacc gatgaacagc tcgcctatgc ctacatcttg 1020
ggcaaagatg gaggaacgcc acttgtctat agtgatgacc tacctgacag cgaagacaaa 1080
gacagtggtc gttgggccga tgtgtggcaa gatccgaaca tgattaacat gcttgccttc 1140
cacaacgcga tgcaaggaca aagcatgact gtagtggcta gcgatcaatg taccttgcta 1200
tttaagcgcg gcaagcaagg cgtggtagga atcaataaat gtggcgagag taagtcggtg 1260
actgtcgata cttaccagca tgagtttaac tggtacaccc cgtaccaaga cgtattgagc 1320
ggcgacatca ccacagtgag ttctcgttat caccaatttg ttttgccagc gcgcagtgca 1380
aggatgtgga aactataa 1398

<210> 176
<211> 465
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 176
Met Lys Asn Ile Ile Arg Leu Cys Ala Ala Ser Ala Ile Leu Thr Val
 1   5   10   15
Ser His Ala Ser Tyr Ala Asp Ala Ile Leu His Ala Phe Asn Trp Gln
            20   25   30
Tyr Thr Asp Val Thr Ala Asn Ala Asn Gln Ile Ala Ala Asn Gly Phe
        35   40   45
Lys Lys Val Leu Ile Ser Pro Ala Met Lys Ser Ser Gly Ser Gln Trp
    50   55   60
Trp Ala Arg Tyr Gln Pro Gln Asp Leu Arg Val Ile Asp Ser Pro Leu
65   70   75   80
Gly Asn Lys Gln Asp Leu Val Ala Met Ile Asn Ala Leu Asn Ser Val
                85   90   95
Gly Val Asp Val Tyr Ala Asp Val Val Leu Asn His Met Ala Asn Glu
            100   105   110
Ser Trp Lys Arg Ser Asp Leu Asn Tyr Pro Gly Ser Glu Val Leu Asn
        115   120   125
Asp Tyr Gln Ser Arg Ser Ala Tyr Tyr Gln Arg Gln Thr Leu Phe Gly
    130   135   140
Asn Leu Gln Glu Asn Leu Phe Ser Glu Asn Asp Phe His Pro Ala Gly
145   150   155   160
Cys Ile Thr Asn Trp Asn Asp Pro Gly His Val Gln Tyr Trp Arg Leu
                165   170   175
Cys Gly Gly Gln Gly Asp Thr Gly Leu Pro Asp Leu Asp Pro Asn Gln
            180   185   190
Trp Val Val Ser Gln Gln Lys Ser Tyr Leu Asn Ala Leu Lys Ser Met
        195   200   205
Gly Ile Lys Gly Phe Arg Ile Asp Ala Val Lys His Met Ser Gln Tyr
    210   215   220
Gln Ile Asp Gln Val Phe Thr Pro Asp Ile Thr Ala Gly Met His Ile
225   230   235   240
Phe Gly Glu Val Ile Thr Ser Gly Gly Gln Gly Asp Ser Gly Tyr Glu
                245   250   255
Ala Phe Leu Ala Pro Tyr Leu Asn Asn Thr Asp His Ala Ala Tyr Asp
            260   265   270
Phe Pro Leu Phe Ala Ser Ile Arg Ala Ala Phe Ser Phe Ser Gly Gly
        275   280   285
Leu Asn Gln Leu His Asn Pro Gln Ala Tyr Gly Gln Ala Leu Gln Asp
    290   295   300
Ser Arg Ala Ile Thr Phe Thr Ile Thr His Asp Ile Pro Thr Asn Asp
305   310   315   320
Gly Phe Arg Tyr Gln Ile Met Asp Pro Thr Asp Glu Gln Leu Ala Tyr
                325   330   335
Ala Tyr Ile Leu Gly Lys Asp Gly Gly Thr Pro Leu Val Tyr Ser Asp
            340   345   350
Asp Leu Pro Asp Ser Glu Asp Lys Asp Ser Gly Arg Trp Ala Asp Val
        355   360   365
Trp Gln Asp Pro Asn Met Ile Asn Met Leu Ala Phe His Asn Ala Met
    370   375   380

Gln Gly Gln Ser Met Thr Val Val Ala Ser Asp Gln Cys Thr Leu Leu
385 390 395 400
Phe Lys Arg Gly Lys Gln Gly Val Val Gly Ile Asn Lys Cys Gly Glu
405 410 415
Ser Lys Ser Val Thr Val Asp Thr Tyr Gln His Glu Phe Asn Trp Tyr
420 425 430
Thr Pro Tyr Gln Asp Val Leu Ser Gly Asp Ile Thr Thr Val Ser Ser
435 440 445
Arg Tyr His Gln Phe Val Leu Pro Ala Arg Ser Ala Arg Met Trp Lys
450 455 460
Leu
465

<210> 177
<211> 1524
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 177
atgaaaacat tcaaattaaa acgcactttt ttaccgctga ccttgctgct cagtgctcct 60
gcctttgctg ggcaaaatgg caccatgatg cagtattttc attggtacgt acctaatgat 120
ggcgcattat ggacgcaggt tgaaagcaat gctccagtac tcgctgaaaa cggttttaca 180
gcgctctggc taccgcccgc atacaaaggc gcgggcggca gtaatgacgt cggttatggc 240
gtctatgata tgtacgattt aggtgagttt gaccaaaaag gctcagtacg aaccaaatac 300
ggcaccaagg ctcagtacat ctctgcaatc aatgccgcgc acaacaacaa tatccaaatt 360
tacggcgacg ttgtgtttaa ccaccgaggt ggcgctgatg ggaagtcgtg ggtcgatacc 420
aagcgcgttg attgggacaa ccgcaatatt gaactgggcg acaaatggat tgaagcttgg 480
gttgagttta attttcctgg ccgcaacgac aaatactcga acttccattg gacttggtat 540
cactttgacg gtgttgactg ggatgatgcc ggcaaagaaa aagcgatctt taaattcaaa 600
ggcgaaggaa aagcatggga ttgggaagtc agctctgaaa aaggcaatta cgactaccta 660
atgtacgccg atttagacat ggatcaccca gaagttaaac aagagctgaa agattggggt 720
gagtggtaca tcaacatgac cggcgttgat ggctttagaa tggatgccgt gaagcacatt 780
aaatatcagt atctacaaga gtggattgat catttacgtt ggaaaacagg caaagagctt 840
ttcaccgttg gtgagtattg gaattacgac gtaaatcaac tgcacaactt tattactaag 900
acctctggca gtatgtcgtt gttcgatgcg ccgcttcaca tgaatttcta caacgcgtca 960
aaatctggcg gcacttacga tatgcgccaa atcatgaatg gcacgttgat gaaggacaac 1020
ccagtcaaag cagtgactct cgtagaaaac cacgatacgc agccattgca ggcgttagag 1080
tcgacagtag attggtggtt caagcctctt gcttacgcat tcattttatt gcgtgaagaa 1140
ggttatccat cggtgttcta cgcagattac tacggcgcgc agtacagcga caaaggttac 1200
aacattaata tggccaaagt gccttacatt gaagaacttg taacactgcg taaagagtat 1260
gcgtatggca aacagaattc ttatctcgac cattgggatg tgattggctg gacccgagag 1320
ggcgatgctg aacatccaaa ctcaatggcg gtgatcatga gtgatggacc gggcggcaca 1380
aaatggatgt ataccggtaa gccaagtacg cgctatgtcg acaagctggg tatccgaact 1440
gaagatgttt ggaccgatgc caatggctgg gcagaatttc ctgtcaatgg tggttcagtc 1500
tcggtttggg tgggcgttaa gtaa 1524

<210> 178
<211> 507
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 178
Met Lys Thr Phe Lys Leu Lys Arg Thr Phe Leu Pro Leu Thr Leu Leu
1 5 10 15
Leu Ser Ala Pro Ala Phe Ala Gly Gln Asn Gly Thr Met Met Gln Tyr
20 25 30
Phe His Trp Tyr Val Pro Asn Asp Gly Ala Leu Trp Thr Gln Val Glu
35 40 45
Ser Asn Ala Pro Val Leu Ala Glu Asn Gly Phe Thr Ala Leu Trp Leu
50 55 60
Pro Pro Ala Tyr Lys Gly Ala Gly Gly Ser Asn Asp Val Gly Tyr Gly
65 70 75 80

```
Val Tyr Asp Met Tyr Asp Leu Gly Glu Phe Asp Gln Lys Gly Ser Val
                85                  90                  95
Arg Thr Lys Tyr Gly Thr Lys Ala Gln Tyr Ile Ser Ala Ile Asn Ala
            100                 105                 110
Ala His Asn Asn Asn Ile Gln Ile Tyr Gly Asp Val Val Phe Asn His
        115                 120                 125
Arg Gly Gly Ala Asp Gly Lys Ser Trp Val Asp Thr Lys Arg Val Asp
    130                 135                 140
Trp Asp Asn Arg Asn Ile Glu Leu Gly Asp Lys Trp Ile Glu Ala Trp
145                 150                 155                 160
Val Glu Phe Asn Phe Pro Gly Arg Asn Asp Lys Tyr Ser Asn Phe His
                165                 170                 175
Trp Thr Trp Tyr His Phe Asp Gly Val Asp Trp Asp Asp Ala Gly Lys
            180                 185                 190
Glu Lys Ala Ile Phe Lys Phe Lys Gly Glu Gly Lys Ala Trp Asp Trp
        195                 200                 205
Glu Val Ser Ser Glu Lys Gly Asn Tyr Asp Tyr Leu Met Tyr Ala Asp
    210                 215                 220
Leu Asp Met Asp His Pro Glu Val Lys Gln Glu Leu Lys Asp Trp Gly
225                 230                 235                 240
Glu Trp Tyr Ile Asn Met Thr Gly Val Asp Gly Phe Arg Met Asp Ala
                245                 250                 255
Val Lys His Ile Lys Tyr Gln Tyr Leu Gln Glu Trp Ile Asp His Leu
            260                 265                 270
Arg Trp Lys Thr Gly Lys Glu Leu Phe Thr Val Gly Glu Tyr Trp Asn
        275                 280                 285
Tyr Asp Val Asn Gln Leu His Asn Phe Ile Thr Lys Thr Ser Gly Ser
    290                 295                 300
Met Ser Leu Phe Asp Ala Pro Leu His Met Asn Phe Tyr Asn Ala Ser
305                 310                 315                 320
Lys Ser Gly Gly Thr Tyr Asp Met Arg Gln Ile Met Asn Gly Thr Leu
                325                 330                 335
Met Lys Asp Asn Pro Val Lys Ala Val Thr Leu Val Glu Asn His Asp
            340                 345                 350
Thr Gln Pro Leu Gln Ala Leu Glu Ser Thr Val Asp Trp Trp Phe Lys
        355                 360                 365
Pro Leu Ala Tyr Ala Phe Ile Leu Leu Arg Glu Glu Gly Tyr Pro Ser
    370                 375                 380
Val Phe Tyr Ala Asp Tyr Tyr Gly Ala Gln Tyr Ser Asp Lys Gly Tyr
385                 390                 395                 400
Asn Ile Asn Met Ala Lys Val Pro Tyr Ile Glu Glu Leu Val Thr Leu
                405                 410                 415
Arg Lys Glu Tyr Ala Tyr Gly Lys Gln Asn Ser Tyr Leu Asp His Trp
            420                 425                 430
Asp Val Ile Gly Trp Thr Arg Glu Gly Asp Ala Glu His Pro Asn Ser
        435                 440                 445
Met Ala Val Ile Met Ser Asp Gly Pro Gly Gly Thr Lys Trp Met Tyr
    450                 455                 460
Thr Gly Lys Pro Ser Thr Arg Tyr Val Asp Lys Leu Gly Ile Arg Thr
465                 470                 475                 480
Glu Asp Val Trp Thr Asp Ala Asn Gly Trp Ala Glu Phe Pro Val Asn
                485                 490                 495
Gly Gly Ser Val Ser Val Trp Val Gly Val Lys
            500                 505
```

<210> 179
<211> 1524
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 179
atgaaaacat tcaaattaaa acgcactttt ttaccgctaa ccttgctgct cagtgctcct 60
gcctttgccg ggcaaaatgg caccatgatg cagtactttc attggtacgt acctaatgat 120
ggcgcattat ggacgcaggt tgaaagcaat gctccagcac tcgctgaaaa cggttttaca 180
gcgctctggc taccgccagc ttacaaaggc gcgggcggca gtaatgatgt cggttatggc 240

```
gtctacgata tgtacgattt aggtgagttt gatcaaaaag gctcagtacg aaccaaatac     300
ggtaccaagg ctcagtacat ctctgcaatc aatgctgcgc acaacaacaa tatccaaatt     360
tacggcgacg ttgtgtttaa ccatcgtggt ggcgctgatg ggaagtcgtg ggtcgatacc     420
aagcgcgttg attgggacaa ccgtaacatt gaactgggcg acaaatggat tgaagcttgg     480
gttgagttta attttcctag ccgcaacgac aaatactcga acttccattg gacttggtat     540
cactttgacg gtgttgactg ggatgatgcc ggcaaagaaa aagcgatctt taaattcaaa     600
ggcgaaggaa aagcatggga ttgggaagtc agctctgaaa aaggcaatta cgactaccta     660
atgtacgccg atttagacat ggatcaccca gaagttaaac aagagctgaa agattggggt     720
gagtggtaca tcaacatgac cggcgttgat ggctttagaa tggatgccgt taagcacatt     780
aaatatcagt atctacaaga gtggattgat catttacgtt ggaaaacagg caaagagctt     840
ttcaccgttg gtgagtattg gaattacgac gtaaatcaac tgcataactt tattactaag     900
acctctggca gtatgtcgtt gttcgatgcg ccgcttcaca tgaacttcta caacgcgtca     960
aaatctggcg gcaattacga tatgcgccaa atcatgaatg gcacgttgat gaaggacaac    1020
ccagtcaaag ctgtgactct cgtagaaaac cacgatacgc agccattgca ggcgttagag    1080
tcgacagtgg attggtggtt caagcctctt gcttacgcat tcatcttgtt gcgtgaagaa    1140
ggttatccat cggtgttcta cgcagattac tacggcgcgc agtacagcga caaaggttac    1200
aacattaata tggccaaagt gccttacatt gaagaacttg taacactgcg taaagagtat    1260
gcgtatggca aacagaattc ttatctcgac cattgggatg tgattggctg gactcgagag    1320
ggcgatgctg aacatccaaa ctcaatggcg gtgatcatga gtgatggacc gggcggaaca    1380
aaatggatgt ataccggtaa tccaagcacg cgctatgtcg acaagctggg tatccgaact    1440
gaagatgttt ggaccgatgc caatggctgg gcagaatttc ctgtcaatgg tggttcagtc    1500
tcggtttggg tgggcgttaa gtaa                                           1524

<210> 180
<211> 507
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 180
Met Lys Thr Phe Lys Leu Lys Arg Thr Phe Leu Pro Leu Thr Leu Leu
 1               5                  10                  15
Leu Ser Ala Pro Ala Phe Ala Gly Gln Asn Gly Thr Met Met Gln Tyr
            20                  25                  30
Phe His Trp Tyr Val Pro Asn Asp Gly Ala Leu Trp Thr Gln Val Glu
        35                  40                  45
Ser Asn Ala Pro Ala Leu Ala Glu Asn Gly Phe Thr Ala Leu Trp Leu
    50                  55                  60
Pro Pro Ala Tyr Lys Gly Ala Gly Gly Ser Asn Asp Val Gly Tyr Gly
65                  70                  75                  80
Val Tyr Asp Met Tyr Asp Leu Gly Glu Phe Asp Gln Lys Gly Ser Val
                85                  90                  95
Arg Thr Lys Tyr Gly Thr Lys Ala Gln Tyr Ile Ser Ala Ile Asn Ala
            100                 105                 110
Ala His Asn Asn Asn Ile Gln Ile Tyr Gly Asp Val Val Phe Asn His
        115                 120                 125
Arg Gly Gly Ala Asp Gly Lys Ser Trp Val Asp Thr Lys Arg Val Asp
    130                 135                 140
Trp Asp Asn Arg Asn Ile Glu Leu Gly Asp Lys Trp Ile Glu Ala Trp
145                 150                 155                 160
Val Glu Phe Asn Phe Pro Ser Arg Asn Asp Lys Tyr Ser Asn Phe His
                165                 170                 175
Trp Thr Trp Tyr His Phe Asp Gly Val Asp Trp Asp Asp Ala Gly Lys
            180                 185                 190
Glu Lys Ala Ile Phe Lys Phe Lys Gly Glu Gly Lys Ala Trp Asp Trp
        195                 200                 205
Glu Val Ser Ser Glu Lys Gly Asn Tyr Asp Tyr Leu Met Tyr Ala Asp
    210                 215                 220
Leu Asp Met Asp His Pro Glu Val Lys Gln Glu Leu Lys Asp Trp Gly
225                 230                 235                 240
Glu Trp Tyr Ile Asn Met Thr Gly Val Asp Gly Phe Arg Met Asp Ala
                245                 250                 255
Val Lys His Ile Lys Tyr Gln Tyr Leu Gln Glu Trp Ile Asp His Leu
            260                 265                 270
Arg Trp Lys Thr Gly Lys Glu Leu Phe Thr Val Gly Glu Tyr Trp Asn
        275                 280                 285
```

```
Tyr Asp Val Asn Gln Leu His Asn Phe Ile Thr Lys Thr Ser Gly Ser
    290                 295                 300
Met Ser Leu Phe Asp Ala Pro Leu His Met Asn Phe Tyr Asn Ala Ser
305                 310                 315                 320
Lys Ser Gly Gly Asn Tyr Asp Met Arg Gln Ile Met Asn Gly Thr Leu
                325                 330                 335
Met Lys Asp Asn Pro Val Lys Ala Val Thr Leu Val Glu Asn His Asp
            340                 345                 350
Thr Gln Pro Leu Gln Ala Leu Glu Ser Thr Val Asp Trp Trp Phe Lys
        355                 360                 365
Pro Leu Ala Tyr Ala Phe Ile Leu Leu Arg Glu Glu Gly Tyr Pro Ser
    370                 375                 380
Val Phe Tyr Ala Asp Tyr Tyr Gly Ala Gln Tyr Ser Asp Lys Gly Tyr
385                 390                 395                 400
Asn Ile Asn Met Ala Lys Val Pro Tyr Ile Glu Glu Leu Val Thr Leu
                405                 410                 415
Arg Lys Glu Tyr Ala Tyr Gly Lys Gln Asn Ser Tyr Leu Asp His Trp
            420                 425                 430
Asp Val Ile Gly Trp Thr Arg Glu Gly Asp Ala Glu His Pro Asn Ser
        435                 440                 445
Met Ala Val Ile Met Ser Asp Gly Pro Gly Gly Thr Lys Trp Met Tyr
    450                 455                 460
Thr Gly Asn Pro Ser Thr Arg Tyr Val Asp Lys Leu Gly Ile Arg Thr
465                 470                 475                 480
Glu Asp Val Trp Thr Asp Ala Asn Gly Trp Ala Glu Phe Pro Val Asn
                485                 490                 495
Gly Gly Ser Val Ser Val Trp Val Gly Val Lys
            500                 505
```

<210> 181
<211> 1830
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 181

```
atgccagagg ccttcggcct ggccattacg ccgtcacata gccggcgggg gaggttggtg      60
ggcgtgtcgc gcgggggcag cctgccgatg ccggtcctcc actggccggc gttcatcctc     120
gtccggcgct tcgtcgccgg tcatccgaac aagcacaaga accggagtat tgcgatgagc     180
cacaccctgc gtgccgccgt attggcggcg atcctgctgc cgttccccgc cctcgctgac     240
caggccggca agagcccggc cggcgtgcgc taccacggcg gcgacgaaat catcctccag     300
ggcttccact ggaacgtcgt ccgcgaagcg cccaacgact ggtacaacat ccttcgccag     360
caggcctcga cgatcgccgc ggacggcttc tcggcaatct ggatgccggt gccctggcgt     420
gacttctcca gctggaccga cggcggcaag tcaggcggcg gcgaaggcta cttctggcac     480
gacttcaaca agaacggccg ctacggcagc gacgcccagc tgcgccaggc cgccggcgca     540
ctcggtggcg ccggggtgaa ggtgctctac gatgtggtgc ccaatcacat gaaccgcggc     600
tatccggaca aggagatcaa cctgccggcc ggccagggct tctggcgcaa cgactgcacc     660
gacccgggca actaccccaa cgactgcgat gacggtgacc gcttcatcgg cggcaagtcg     720
gacctgaaca ccggccatcc gcagatctac ggcatgtttc gcgacgagct tgccaacctg     780
cgcagcgggt acggcgccgg cggcttccgc ttcgacttcg ttcgcggcta tgcgcccgaa     840
cgggtcgaca gctggatgag cgacagcgcc gacagcagtt tctgcgttgg cgagctgtgg     900
aaaagcccgt ccgagtaccc gagctgggac tggcgcaaca cggcgagctg gcagcagatc     960
atcaaggact ggtccgaccg ggccaagtgc ccggtgttcg acttcgcgct caaggagcgc    1020
atgcagaacg gctcggtcgc cgactggaag catggcctca atggcaaccc ggacccgcgc    1080
tggcgcgagg tggcggtgac ctttgtcgac aaccacgaca ccggctattc gcccgggcag    1140
aacggcggcc agcaccactg ggcgctgcag gacgggctga tccgccaggc ctacgcctac    1200
atcctcacca gcccgggcac gccggtggtg tactggtcgc acatgtacga ctggggctac    1260
ggcgacttca ttcgccagct gatccaggtg cggcgcaccg ctggcgtgcg cgccgattcg    1320
gcgatcagct tccacagcgg ctacagcggc ctggtcgcta ccgtcagcgg cagccatcag    1380
accctggtgg tggcgctcaa ctccgatctg gccaaccccg gccaggtcgc cagcggcagc    1440
ttcagcgagg cggtcaacgc cagcaacggc caggtgcgcg tctggcgcag cggtagcggc    1500
gatggcggcg gcaatgacgg cggcgagggc ggtctggtca atgtgaactt ccgctgcgac    1560
aacggcgtga cgcagatggg cgacagcgtc tacgcggtgg gcaacgtcag ccagctcggc    1620
aactggagcc cggcctccgc ggtacggctg accgacacca gcagctatcc gacctggaag    1680
ggcagcatcg ccctgcctga cggtcagaac gtggaatgga agtgcctgat ccgtaacgag    1740
gcggacgcga cgctggtgcg ccagtggcaa tcgggcggca acaaccaggt ccaggccgct    1800
```

gccggcgcga gcaccagcgg ctcgttctga 1830

<210> 182
<211> 609
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 182

```
Met Pro Glu Ala Phe Gly Leu Ala Ile Thr Pro Ser His Ser Arg Arg
 1               5                  10                  15
Gly Arg Leu Val Gly Val Ser Arg Gly Gly Ser Leu Pro Met Pro Val
            20                  25                  30
Leu His Trp Pro Ala Phe Ile Leu Val Arg Arg Phe Val Ala Gly His
        35                  40                  45
Pro Asn Lys His Lys Asn Arg Ser Ile Ala Met Ser His Thr Leu Arg
    50                  55                  60
Ala Ala Val Leu Ala Ala Ile Leu Leu Pro Phe Pro Ala Leu Ala Asp
65                  70                  75                  80
Gln Ala Gly Lys Ser Pro Ala Gly Val Arg Tyr His Gly Gly Asp Glu
                85                  90                  95
Ile Ile Leu Gln Gly Phe His Trp Asn Val Val Arg Glu Ala Pro Asn
            100                 105                 110
Asp Trp Tyr Asn Ile Leu Arg Gln Gln Ala Ser Thr Ile Ala Ala Asp
        115                 120                 125
Gly Phe Ser Ala Ile Trp Met Pro Val Pro Trp Arg Asp Phe Ser Ser
    130                 135                 140
Trp Thr Asp Gly Gly Lys Ser Gly Gly Gly Glu Gly Tyr Phe Trp His
145                 150                 155                 160
Asp Phe Asn Lys Asn Gly Arg Tyr Gly Ser Asp Ala Gln Leu Arg Gln
                165                 170                 175
Ala Ala Gly Ala Leu Gly Gly Ala Gly Val Lys Val Leu Tyr Asp Val
            180                 185                 190
Val Pro Asn His Met Asn Arg Gly Tyr Pro Asp Lys Glu Ile Asn Leu
        195                 200                 205
Pro Ala Gly Gln Gly Phe Trp Arg Asn Asp Cys Thr Asp Pro Gly Asn
    210                 215                 220
Tyr Pro Asn Asp Cys Asp Asp Gly Asp Arg Phe Ile Gly Gly Lys Ser
225                 230                 235                 240
Asp Leu Asn Thr Gly His Pro Gln Ile Tyr Gly Met Phe Arg Asp Glu
                245                 250                 255
Leu Ala Asn Leu Arg Ser Gly Tyr Gly Ala Gly Gly Phe Arg Phe Asp
            260                 265                 270
Phe Val Arg Gly Tyr Ala Pro Glu Arg Val Asp Ser Trp Met Ser Asp
        275                 280                 285
Ser Ala Asp Ser Ser Phe Cys Val Gly Glu Leu Trp Lys Ser Pro Ser
    290                 295                 300
Glu Tyr Pro Ser Trp Asp Trp Arg Asn Thr Ala Ser Trp Gln Gln Ile
305                 310                 315                 320
Ile Lys Asp Trp Ser Asp Arg Ala Lys Cys Pro Val Phe Asp Phe Ala
                325                 330                 335
Leu Lys Glu Arg Met Gln Asn Gly Ser Val Ala Asp Trp Lys His Gly
            340                 345                 350
Leu Asn Gly Asn Pro Asp Pro Arg Trp Arg Glu Val Ala Val Thr Phe
        355                 360                 365
Val Asp Asn His Asp Thr Gly Tyr Ser Pro Gly Gln Asn Gly Gly Gln
    370                 375                 380
His His Trp Ala Leu Gln Asp Gly Leu Ile Arg Gln Ala Tyr Ala Tyr
385                 390                 395                 400
Ile Leu Thr Ser Pro Gly Thr Pro Val Val Tyr Trp Ser His Met Tyr
                405                 410                 415
Asp Trp Gly Tyr Gly Asp Phe Ile Arg Gln Leu Ile Gln Val Arg Arg
            420                 425                 430
Thr Ala Gly Val Arg Ala Asp Ser Ala Ile Ser Phe His Ser Gly Tyr
        435                 440                 445
Ser Gly Leu Val Ala Thr Val Ser Gly Ser His Gln Thr Leu Val Val
```

```
    450                 455                 460
Ala Leu Asn Ser Asp Leu Ala Asn Pro Gly Gln Val Ala Ser Gly Ser
465                 470                 475                 480
Phe Ser Glu Ala Val Asn Ala Ser Asn Gly Gln Val Arg Val Trp Arg
                485                 490                 495
Ser Gly Ser Gly Asp Gly Gly Gly Asn Asp Gly Gly Glu Gly Gly Leu
            500                 505                 510
Val Asn Val Asn Phe Arg Cys Asp Asn Gly Val Thr Gln Met Gly Asp
        515                 520                 525
Ser Val Tyr Ala Val Gly Asn Val Ser Gln Leu Gly Asn Trp Ser Pro
    530                 535                 540
Ala Ser Ala Val Arg Leu Thr Asp Thr Ser Ser Tyr Pro Thr Trp Lys
545                 550                 555                 560
Gly Ser Ile Ala Leu Pro Asp Gly Gln Asn Val Glu Trp Lys Cys Leu
                565                 570                 575
Ile Arg Asn Glu Ala Asp Ala Thr Leu Val Arg Gln Trp Gln Ser Gly
            580                 585                 590
Gly Asn Asn Gln Val Gln Ala Ala Ala Gly Ala Ser Thr Ser Gly Ser
        595                 600                 605
Phe
```

<210> 183
<211> 1596
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 183

```
atgcaaacga ttgcaaaaaa aggggatgaa acgatgaaag ggaaaaaatg gacagcttta      60
gctctaacac tgccgctggc tgctagctta tcaacaggcg ttcacgccga aaccgtacat     120
aaaggtaagt ctgaagcaac agataaaaac ggtgtctttt atgaggtgta tgtaaactct     180
ttttacgata caaataaaga tggacatggt gatttaaaag gtctgacaca aaagttggat     240
tatttaaatg acggcaattc tcatacaaag aatgatcttc aagtaaacgg gatttggatg     300
atgccagtca acccttctcc tagctatcat aaatatgatg taacggacta ttataacatt     360
gatcctcagt acggaaatct gcaagatttt cgcaagctga tgaaagaagc agacaaacga     420
gacgtaaaag tcattatgga ccttgttgtg aatcatacga gcagcgaaca cccttggttt     480
caagctgcat taaaagataa aaacagcaag tacagagatt actatatttg ggctgataaa     540
aataccgatt tgaatgaaaa aggatcttgg gggcagcaag tatggcataa agctccaaac     600
ggagagtatt tttacggaac gttttgggaa ggaatgcctg acttaaatta cgataaccct     660
gaagtaagaa aagaaatgat taacgtcgga aagttttggc taaagcaagg cgttaatggc     720
ttccgcttag atgctgcgct tcatattttt aaaggtcaaa cacctgaagg cgctaagaaa     780
aatatcctgt ggtggaatga gtttagagat gcgatgaaaa aagaaaaccc taacgtatat     840
ctaacgggtg aagtatggga tcagcctgaa gtggtagctc cttactatca atcgcttgat     900
tctttattta attttgattt agcaggaaaa attgtcagct ctgtaaaagc aggaaatgat     960
caaggaatcg ccactgcagc agcggcaaca gatgaactgt tcaaatcata caatccaaat    1020
aaaattgacg gcattttctt aaccaaccat gaccaaaatc gcgtcatgag tgagctgagc    1080
ggcgatgtga acaaagcaaa atcagctgct tctatcttac ttacgcttcc tggcaacccg    1140
tatatttatt acggtgaaga aattggcatg accggtgaaa agcctgatga gttaatccgt    1200
gaaccattcc gctggtacga aggaaacgga cttggacaaa ctagctggga aacacctgta    1260
tataacaaag gcggcaacgg cgtgtctgta gaagtacaaa ccaaacaaaa ggattctttg    1320
ttaaatcatt atcgtgaaat gattcgcgtg cgtcagcagc atgaagagtt agtaaaagga    1380
acgcttcaat ctatttcagt agacagtaaa gaagtggttg cctatagtcg cacgtataaa    1440
ggcaactcga ttagcgtgta tcataatatt tcaaatcaac ctgtaaaagt atctgtagca    1500
gcgaaaggta aattgatttt tgctagtgaa aaaggtgcta aaaaagtcaa aaatcagctt    1560
gtaattccgg ctaatacaac ggttttaata aaataa                              1596
```

<210> 184
<211> 531
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 184

```
Met Gln Thr Ile Ala Lys Lys Gly Asp Glu Thr Met Lys Gly Lys Lys
 1               5                  10                  15
Trp Thr Ala Leu Ala Leu Thr Leu Pro Leu Ala Ala Ser Leu Ser Thr
            20                  25                  30
Gly Val His Ala Glu Thr Val His Lys Gly Lys Ser Glu Ala Thr Asp
        35                  40                  45
Lys Asn Gly Val Phe Tyr Glu Val Tyr Val Asn Ser Phe Tyr Asp Thr
    50                  55                  60
Asn Lys Asp Gly His Gly Asp Leu Lys Gly Leu Thr Gln Lys Leu Asp
65                  70                  75                  80
Tyr Leu Asn Asp Gly Asn Ser His Thr Lys Asn Asp Leu Gln Val Asn
                85                  90                  95
Gly Ile Trp Met Met Pro Val Asn Pro Ser Pro Ser Tyr His Lys Tyr
            100                 105                 110
Asp Val Thr Asp Tyr Tyr Asn Ile Asp Pro Gln Tyr Gly Asn Leu Gln
        115                 120                 125
Asp Phe Arg Lys Leu Met Lys Glu Ala Asp Lys Arg Asp Val Lys Val
    130                 135                 140
Ile Met Asp Leu Val Val Asn His Thr Ser Ser Glu His Pro Trp Phe
145                 150                 155                 160
Gln Ala Ala Leu Lys Asp Lys Asn Ser Lys Tyr Arg Asp Tyr Tyr Ile
                165                 170                 175
Trp Ala Asp Lys Asn Thr Asp Leu Asn Glu Lys Gly Ser Trp Gly Gln
            180                 185                 190
Gln Val Trp His Lys Ala Pro Asn Gly Glu Tyr Phe Tyr Gly Thr Phe
        195                 200                 205
Trp Glu Gly Met Pro Asp Leu Asn Tyr Asp Asn Pro Glu Val Arg Lys
    210                 215                 220
Glu Met Ile Asn Val Gly Lys Phe Trp Leu Lys Gln Gly Val Asn Gly
225                 230                 235                 240
Phe Arg Leu Asp Ala Ala Leu His Ile Phe Lys Gly Gln Thr Pro Glu
                245                 250                 255
Gly Ala Lys Lys Asn Ile Leu Trp Trp Asn Glu Phe Arg Asp Ala Met
            260                 265                 270
Lys Lys Glu Asn Pro Asn Val Tyr Leu Thr Gly Glu Val Trp Asp Gln
        275                 280                 285
Pro Glu Val Val Ala Pro Tyr Tyr Gln Ser Leu Asp Ser Leu Phe Asn
    290                 295                 300
Phe Asp Leu Ala Gly Lys Ile Val Ser Ser Val Lys Ala Gly Asn Asp
305                 310                 315                 320
Gln Gly Ile Ala Thr Ala Ala Ala Ala Thr Asp Glu Leu Phe Lys Ser
                325                 330                 335
Tyr Asn Pro Asn Lys Ile Asp Gly Ile Phe Leu Thr Asn His Asp Gln
            340                 345                 350
Asn Arg Val Met Ser Glu Leu Ser Gly Asp Val Asn Lys Ala Lys Ser
        355                 360                 365
Ala Ala Ser Ile Leu Leu Thr Leu Pro Gly Asn Pro Tyr Ile Tyr Tyr
    370                 375                 380
Gly Glu Glu Ile Gly Met Thr Gly Glu Lys Pro Asp Glu Leu Ile Arg
385                 390                 395                 400
Glu Pro Phe Arg Trp Tyr Glu Gly Asn Gly Leu Gly Gln Thr Ser Trp
                405                 410                 415
Glu Thr Pro Val Tyr Asn Lys Gly Gly Asn Gly Val Ser Val Glu Val
            420                 425                 430
Gln Thr Lys Gln Lys Asp Ser Leu Leu Asn His Tyr Arg Glu Met Ile
        435                 440                 445
Arg Val Arg Gln Gln His Glu Glu Leu Val Lys Gly Thr Leu Gln Ser
    450                 455                 460
Ile Ser Val Asp Ser Lys Glu Val Val Ala Tyr Ser Arg Thr Tyr Lys
465                 470                 475                 480
Gly Asn Ser Ile Ser Val Tyr His Asn Ile Ser Asn Gln Pro Val Lys
                485                 490                 495
Val Ser Val Ala Ala Lys Gly Lys Leu Ile Phe Ala Ser Glu Lys Gly
            500                 505                 510
Ala Lys Lys Val Lys Asn Gln Leu Val Ile Pro Ala Asn Thr Thr Val
        515                 520                 525
Leu Ile Lys
    530
```

<210> 185
<211> 1572
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 185
atgaaactga tgaaagggaa aaaatggaca gctttagctc taacactgcc gctggctgct 60
agcttatcaa caggcgttca cgccgaaact gtacataaag gtaaagctcc aacagcagat 120
aaaaacggtg tcttttatga ggtgtatgta aactcttttt acgatgcaaa taaagatgga 180
catggtgatt taaaaggtct tacacaaaag ctggactatt taaatgacgg aaattctcat 240
acaaagaatg atcttcaagt aaacgggatt tggatgatgc cagtcaaccc ttctcctagc 300
tatcataaat atgatgtaac ggattattat aacattgatc cgcagtacgg aaatctgcaa 360
gattttcgca agctgatgaa agaagcagac aaacgagacg taaaagtcat tatggacctt 420
gttgtgaatc atacgagcag cgaacaccct tggtttcaag ctgcgttaaa agataaaaac 480
agcaagtaca gagattacta tatttgggct gataaaaata ccgacttgaa tgaaaaagga 540
tcttggggac agcaagtatg gcataaagct ccaaacggag agtattttta cggaacgttt 600
tgggaaggaa tgcctgactt aaattacgat aaccctgaag taagaaaaga aatgattaac 660
gtcggaaagt tttggctaaa gcaaggcgtt gatggcttcc gcttagatgc tgcgcttcat 720
atttttaaag gtcaaacgcc tgaaggcgct aagaaaaata ttctgtggtg gaatgagttt 780
agagatgcga tgaaaaaaga aaaccctaac gtatatctaa cgggtgaagt atgggatcag 840
cctgaagtgg tagctcctta ctatcaatcg cttgattccc tatttaactt tgatttagca 900
gggaaaattg tcagttctgt aaaagcagga aatgatcaag gaatcgccac tgcagcagcg 960
gcaacggatg agctgttcaa atcatacaat ccaaataaaa ttgacggcat tttcttaacc 1020
aaccatgacc aaaaccgcgt catgagtgaa ctgatcggcg atgtgaacaa agcaaaatca 1080
gctgcttcta tcttacttac gcttcctggc aacccgtata tttattacgg tgaagaaatt 1140
ggcatgaccg gtgaaaagcc tgatgagtta atccgtgaac cgttccgctg gtacgaagga 1200
aacggacttg gacaaaccag ctgggaaaca cctgtatata acaaaggcgg caacggcgtg 1260
tctgtagaag cacaaaccaa acaaaaggat tctttgttaa atcattaccg tgaaatgatt 1320
cgcgtgcgtc agcagcatga agagttagta aaaggaacgc ttcaatctat tttagtagac 1380
agtaaagaag ttgttgccta tagccgtacg tataaagaca actcgattag cgtgtatcat 1440
aatatttcaa atcaaccggt aaaagtatct gtagcagcaa aaggtaaatt aatttttgct 1500
agtgaaaaag gtgctaaaaa agtcaagaat cagcttgtga ttccggctaa tacaacggtt 1560
ttaataaaat aa 1572

<210> 186
<211> 523
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 186
Met Lys Leu Met Lys Gly Lys Lys Trp Thr Ala Leu Ala Leu Thr Leu
1 5 10 15
Pro Leu Ala Ala Ser Leu Ser Thr Gly Val His Ala Glu Thr Val His
20 25 30
Lys Gly Lys Ala Pro Thr Ala Asp Lys Asn Gly Val Phe Tyr Glu Val
35 40 45
Tyr Val Asn Ser Phe Tyr Asp Ala Asn Lys Asp Gly His Gly Asp Leu
50 55 60
Lys Gly Leu Thr Gln Lys Leu Asp Tyr Leu Asn Asp Gly Asn Ser His
65 70 75 80
Thr Lys Asn Asp Leu Gln Val Asn Gly Ile Trp Met Met Pro Val Asn
85 90 95
Pro Ser Pro Ser Tyr His Lys Tyr Asp Val Thr Asp Tyr Tyr Asn Ile
100 105 110
Asp Pro Gln Tyr Gly Asn Leu Gln Asp Phe Arg Lys Leu Met Lys Glu
115 120 125
Ala Asp Lys Arg Asp Val Lys Val Ile Met Asp Leu Val Val Asn His
130 135 140
Thr Ser Ser Glu His Pro Trp Phe Gln Ala Ala Leu Lys Asp Lys Asn
145 150 155 160
Ser Lys Tyr Arg Asp Tyr Tyr Ile Trp Ala Asp Lys Asn Thr Asp Leu

```
                165                 170                 175
Asn Glu Lys Gly Ser Trp Gly Gln Gln Val Trp His Lys Ala Pro Asn
            180                 185                 190
Gly Glu Tyr Phe Tyr Gly Thr Phe Trp Glu Gly Met Pro Asp Leu Asn
        195                 200                 205
Tyr Asp Asn Pro Glu Val Arg Lys Glu Met Ile Asn Val Gly Lys Phe
    210                 215                 220
Trp Leu Lys Gln Gly Val Asp Gly Phe Arg Leu Asp Ala Ala Leu His
225                 230                 235                 240
Ile Phe Lys Gly Gln Thr Pro Glu Gly Ala Lys Lys Asn Ile Leu Trp
                245                 250                 255
Trp Asn Glu Phe Arg Asp Ala Met Lys Lys Glu Asn Pro Asn Val Tyr
            260                 265                 270
Leu Thr Gly Glu Val Trp Asp Gln Pro Glu Val Val Ala Pro Tyr Tyr
        275                 280                 285
Gln Ser Leu Asp Ser Leu Phe Asn Phe Asp Leu Ala Gly Lys Ile Val
    290                 295                 300
Ser Ser Val Lys Ala Gly Asn Asp Gln Gly Ile Ala Thr Ala Ala Ala
305                 310                 315                 320
Ala Thr Asp Glu Leu Phe Lys Ser Tyr Asn Pro Asn Lys Ile Asp Gly
                325                 330                 335
Ile Phe Leu Thr Asn His Asp Gln Asn Arg Val Met Ser Glu Leu Ile
            340                 345                 350
Gly Asp Val Asn Lys Ala Lys Ser Ala Ala Ser Ile Leu Leu Thr Leu
        355                 360                 365
Pro Gly Asn Pro Tyr Ile Tyr Tyr Gly Glu Glu Ile Gly Met Thr Gly
    370                 375                 380
Glu Lys Pro Asp Glu Leu Ile Arg Glu Pro Phe Arg Trp Tyr Glu Gly
385                 390                 395                 400
Asn Gly Leu Gly Gln Thr Ser Trp Glu Thr Pro Val Tyr Asn Lys Gly
                405                 410                 415
Gly Asn Gly Val Ser Val Glu Ala Gln Thr Lys Gln Lys Asp Ser Leu
            420                 425                 430
Leu Asn His Tyr Arg Glu Met Ile Arg Val Arg Gln Gln His Glu Glu
        435                 440                 445
Leu Val Lys Gly Thr Leu Gln Ser Ile Leu Val Asp Ser Lys Glu Val
    450                 455                 460
Val Ala Tyr Ser Arg Thr Tyr Lys Asp Asn Ser Ile Ser Val Tyr His
465                 470                 475                 480
Asn Ile Ser Asn Gln Pro Val Lys Val Ser Val Ala Ala Lys Gly Lys
                485                 490                 495
Leu Ile Phe Ala Ser Glu Lys Gly Ala Lys Lys Val Lys Asn Gln Leu
            500                 505                 510
Val Ile Pro Ala Asn Thr Thr Val Leu Ile Lys
        515                 520
```

<210> 187
<211> 2052
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 187

```
atgtatctca tccaggaggg gcacatgcgt tttccgccca ttattcaccc gcttaccggc      60
ctggccgttc cggttggagc tctgcgtacc gcacagagct gcggcatagg ggagtttgcc     120
gacttgccgg ttcttgccga attctgcaaa aaagccggat ttgatcttgt acagcttctt     180
ccggtcaatg acaccggcac agaaagttct ccatacagcg cgctttctgc ctttgccctg     240
cacccgctgt atatcaggct ttccgacctg cctgaagcag cgggtttcga aaagcagatt     300
acagatctga aaagccggtt tgaggacttg cctcgtttca gctatacgga gctgcgccgt     360
gccaaactgg atatcctgcg tgcagtgttt gataaaaaca aggcaaccat catcggcagt     420
gccgaactgg aagcctggat ttcagataac ccctggatca tcgaatatgc ggtttttatg     480
aaccagaaac accgcaactt tgaagccggc tggaaacatt gggaaaagct gcgcaacccc     540
actcataacg aaatacaaaa aacctggcag ggtaaaacct ggcaggctga ccatcaattc     600
tttgcatggc tgcagatgcg gctggaccag cagtttactg ccgccgctac agagtgcaac     660
gccctgggtg tctatcttaa gggcgatata cctataatga tgaacgagga ttccgcagat     720
gcctgggcga atccggaatt cttccgtgac gatcttcggg ccggaagtcc ccctgacggt     780
```

gaaaaccccc agggacaaaa ctggggcttc cccatttata actgggaaaa ccttgcaaat 840
gacgggtaca gctggtggaa aaaacgtctg aagcacagcg cacggtatta ccatgcctac 900
cgcattgacc atattcttgg gtttttccgg atatgggcta taccctatgg cgaatactcc 960
ggctacctgg gatggccctt gccgcatgaa ccggtaagtg cagcagaact ggcagaacgg 1020
ggcttttcca aggaccgctt gcgctggctt accgaacccc acttgcctac acgggcagcc 1080
gaggaagcga ataactggga ctatctggga acacacggct atctgaatca gatcatgaac 1140
cgtatcggtg aagaagaact atggctgttc aagcccgaga tcacctgcga ggcagatata 1200
cgaaacacaa acctgccgga tgccctgaaa gaggttctgg tacggcagtg gaaaaaccgg 1260
ctgctgcagg ttaccggccg cgacgaaaaa ggacggacaa tctactatcc gctgtggcgt 1320
ttccgtgaca gcactgcatg gcagacgctt accgatggcg agaaacactc cctggaagag 1380
ctgttcgccc aaaaagcggc gcacaatgaa accctgtggc gagaacaggc ggtggaactt 1440
ctgggtgagc tgacgcgatc tacggatatg cttgcctgtg ctgaagatct gggaagtatt 1500
ccccacagtg taccggaagt gctttcaaac ctttcaattt acagtctgcg ggttacccgc 1560
tgggcccgcc aatgggatgc ccccggccag ccctttcaca gactggagga gtatccgctc 1620
atgtcggtag cgaccccatc ggttcatgat tcctctaccc tgcgcggatg gtgggaaacc 1680
gaaggcggcg accgggcctt tatggacgca tggcctccgg aacaggatgc atacgcagga 1740
gcaggccgcc atgagttcga aggcgcctgg ggaccccgcc aggcatcctg ggtactccgt 1800
aaactctgcg aagcccgttc cgcgctctgt gttttcccca tccaggatat tttggccctg 1860
tcttcagact tttatgcaat gacagcggac gaggaacgca tcaatattcc gggcagtgta 1920
tccggattta actggacata ccggttgcct gcggcaatcg aggatttatc taaaaacagc 1980
caacttataa ccgcaatcca gaccgcgttg caggaccgcc gggcgaggaa ggcacaagga 2040
gcacagcaat ga 2052

<210> 188
<211> 683
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 188
Met Tyr Leu Ile Gln Glu Gly His Met Arg Phe Pro Pro Ile Ile His
 1   5   10   15
Pro Leu Thr Gly Leu Ala Val Pro Val Gly Ala Leu Arg Thr Ala Gln
            20   25   30
Ser Cys Gly Ile Gly Glu Phe Ala Asp Leu Pro Val Leu Ala Glu Phe
        35   40   45
Cys Lys Lys Ala Gly Phe Asp Leu Val Gln Leu Leu Pro Val Asn Asp
    50   55   60
Thr Gly Thr Glu Ser Ser Pro Tyr Ser Ala Leu Ser Ala Phe Ala Leu
65   70   75   80
His Pro Leu Tyr Ile Arg Leu Ser Asp Leu Pro Glu Ala Ala Gly Phe
                85   90   95
Glu Lys Gln Ile Thr Asp Leu Lys Ser Arg Phe Glu Asp Leu Pro Arg
            100   105   110
Phe Ser Tyr Thr Glu Leu Arg Arg Ala Lys Leu Asp Ile Leu Arg Ala
        115   120   125
Val Phe Asp Lys Asn Lys Ala Thr Ile Ile Gly Ser Ala Glu Leu Glu
    130   135   140
Ala Trp Ile Ser Asp Asn Pro Trp Ile Ile Glu Tyr Ala Val Phe Met
145   150   155   160
Asn Gln Lys His Arg Asn Phe Glu Ala Gly Trp Lys His Trp Glu Lys
                165   170   175
Leu Arg Asn Pro Thr His Asn Glu Ile Gln Lys Thr Trp Gln Gly Lys
            180   185   190
Thr Trp Gln Ala Asp His Gln Phe Phe Ala Trp Leu Gln Met Arg Leu
        195   200   205
Asp Gln Gln Phe Thr Ala Ala Ala Thr Glu Cys Asn Ala Leu Gly Val
    210   215   220
Tyr Leu Lys Gly Asp Ile Pro Ile Met Met Asn Glu Asp Ser Ala Asp
225   230   235   240
Ala Trp Ala Asn Pro Glu Phe Phe Arg Asp Asp Leu Arg Ala Gly Ser
                245   250   255
Pro Pro Asp Gly Glu Asn Pro Gln Gly Gln Asn Trp Gly Phe Pro Ile
            260   265   270
Tyr Asn Trp Glu Asn Leu Ala Asn Asp Gly Tyr Ser Trp Trp Lys Lys
        275   280   285

```
Arg Leu Lys His Ser Ala Arg Tyr Tyr His Ala Tyr Arg Ile Asp His
    290                 295                 300
Ile Leu Gly Phe Phe Arg Ile Trp Ala Ile Pro Tyr Gly Glu Tyr Ser
305                 310                 315                 320
Gly Tyr Leu Gly Trp Pro Leu Pro His Glu Pro Val Ser Ala Ala Glu
                325                 330                 335
Leu Ala Glu Arg Gly Phe Ser Lys Asp Arg Leu Arg Trp Leu Thr Glu
            340                 345                 350
Pro His Leu Pro Thr Arg Ala Ala Glu Glu Ala Asn Asn Trp Asp Tyr
        355                 360                 365
Leu Gly Thr His Gly Tyr Leu Asn Gln Ile Met Asn Arg Ile Gly Glu
    370                 375                 380
Glu Glu Leu Trp Leu Phe Lys Pro Glu Ile Thr Cys Glu Ala Asp Ile
385                 390                 395                 400
Arg Asn Thr Asn Leu Pro Asp Ala Leu Lys Glu Val Leu Val Arg Gln
                405                 410                 415
Trp Lys Asn Arg Leu Leu Gln Val Thr Gly Arg Asp Glu Lys Gly Arg
            420                 425                 430
Thr Ile Tyr Tyr Pro Leu Trp Arg Phe Arg Asp Ser Thr Ala Trp Gln
        435                 440                 445
Thr Leu Thr Asp Gly Glu Lys His Ser Leu Glu Glu Leu Phe Ala Gln
    450                 455                 460
Lys Ala Ala His Asn Glu Thr Leu Trp Arg Glu Gln Ala Val Glu Leu
465                 470                 475                 480
Leu Gly Glu Leu Thr Arg Ser Thr Asp Met Leu Ala Cys Ala Glu Asp
                485                 490                 495
Leu Gly Ser Ile Pro His Ser Val Pro Glu Val Leu Ser Asn Leu Ser
            500                 505                 510
Ile Tyr Ser Leu Arg Val Thr Arg Trp Ala Arg Gln Trp Asp Ala Pro
        515                 520                 525
Gly Gln Pro Phe His Arg Leu Glu Glu Tyr Pro Leu Met Ser Val Ala
    530                 535                 540
Thr Pro Ser Val His Asp Ser Ser Thr Leu Arg Gly Trp Trp Glu Thr
545                 550                 555                 560
Glu Gly Gly Asp Arg Ala Phe Met Asp Ala Trp Pro Pro Glu Gln Asp
                565                 570                 575
Ala Tyr Ala Gly Ala Gly Arg His Glu Phe Glu Gly Ala Trp Gly Pro
            580                 585                 590
Arg Gln Ala Ser Trp Val Leu Arg Lys Leu Cys Glu Ala Arg Ser Ala
        595                 600                 605
Leu Cys Val Phe Pro Ile Gln Asp Ile Leu Ala Leu Ser Ser Asp Phe
    610                 615                 620
Tyr Ala Met Thr Ala Asp Glu Glu Arg Ile Asn Ile Pro Gly Ser Val
625                 630                 635                 640
Ser Gly Phe Asn Trp Thr Tyr Arg Leu Pro Ala Ala Ile Glu Asp Leu
                645                 650                 655
Ser Lys Asn Ser Gln Leu Ile Thr Ala Ile Gln Thr Ala Leu Gln Asp
            660                 665                 670
Arg Arg Ala Arg Lys Ala Gln Gly Ala Gln Gln
        675                 680
```

<210> 189
<211> 1596
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 189

```
atgcaaacga ttgcaaaaaa aggggatgaa acgatgaaag ggaaaaaatg gacagcttta      60
gctctaacac tgccgctggc tgctagctta tcaacaggcg ttcacgccga aaccgtacat     120
aaaggtaaat ctccagctgc agataaaaac ggtgtctttt atgaggtgta tgtaaactct     180
ttttacgatg caaataaaga tggacatggt gatttaaaag gtcttacaca aaaactggac     240
tatttaaatg atggcaattc tcatacaaag aatgatcttc aagtaaacgg gatttggatg     300
atgccgatca acccttctcc tagctatcat aaatatgatg taacggacta ttataacatt     360
gattctcagt acggaaatct gcaagatttt cgcaagctaa tgaaagaagc agataaacga     420
gatgtaaaag ttattatgga cctcgttgtg aatcatacga gcagtgaaca cccttggttt     480
```

caagctgcgt taaaagataa aaacagcaag tacagagatt actatatttg ggctgataaa 540
aataccgatt tgaatgaaaa aggatcttgg ggacaacaag tatggcacaa agctccaaac 600
ggagagtatt tttacggaac gttctgggaa ggaatgcctg acttaaatta cgataaccct 660
gaagtaagaa aagaaatgat taacgtcgga aagttttggc taaagcaagg cgttgacggc 720
ttccgcttag atgctgccct tcatatcttt aaaggtcaaa cacctgaagg cgctaagaaa 780
aatattgtgt ggtggaatga atttagagat gcgatgaaaa aagaaaaccc gaacgtatat 840
ctaacgggcg aagtatggga tcagccggaa gtggtagctc cttattatca gtcgcttgat 900
tccctattta actttgattt agcaggaaaa attgtcagct ctgtaaaagc aggaaatgat 960
caaggaatcg ctactgcagc agcggcaaca gatgaactgt tcaaatcata caatccaaat 1020
aaaattgacg gcattttctt aaccaatcat gaccaaaatc gcgtcatgag tgagttaagc 1080
ggagatgtca ataaagcaaa gtcagctgcc tctatcttac ttacgcttcc tggaaatccg 1140
tatatttatt acggtgaaga aatcggcatg accggtgaaa agcctgatga attaatccgt 1200
gaaccgttcc gctggtacga aggaaacgga cttggacaaa ctagttggga aacacctgta 1260
tacaataaag gcggcaacgg cgtgtctgta gaagcacaaa ccaaacaaaa ggactctttg 1320
ttaaatcatt accgtgaaat gattcgcgtg cgtcagcagc acgaagagtt agtaaaagga 1380
acgcttcaat ctatttcagt agacagtaaa gaagttgttg cttatagccg tacgtataaa 1440
ggcaactcca ttagtgtgta tcataatatt tcaaatcaac ctgtaaaagt atctgtagca 1500
gcgaaaggta aattgatttt tgctagtgaa aaaggtgcta aaaaggtcaa aaatcagctt 1560
gtgattccgg cgaatacaac ggttttagta aaataa 1596

<210> 190
<211> 531
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 190
Met Gln Thr Ile Ala Lys Lys Gly Asp Glu Thr Met Lys Gly Lys Lys
 1   5   10   15
Trp Thr Ala Leu Ala Leu Thr Leu Pro Leu Ala Ala Ser Leu Ser Thr
            20   25   30
Gly Val His Ala Glu Thr Val His Lys Gly Lys Ser Pro Ala Ala Asp
        35   40   45
Lys Asn Gly Val Phe Tyr Glu Val Tyr Val Asn Ser Phe Tyr Asp Ala
    50   55   60
Asn Lys Asp Gly His Gly Asp Leu Lys Gly Leu Thr Gln Lys Leu Asp
65   70   75   80
Tyr Leu Asn Asp Gly Asn Ser His Thr Lys Asn Asp Leu Gln Val Asn
                85   90   95
Gly Ile Trp Met Met Pro Ile Asn Pro Ser Pro Ser Tyr His Lys Tyr
            100   105   110
Asp Val Thr Asp Tyr Tyr Asn Ile Asp Ser Gln Tyr Gly Asn Leu Gln
        115   120   125
Asp Phe Arg Lys Leu Met Lys Glu Ala Asp Lys Arg Asp Val Lys Val
    130   135   140
Ile Met Asp Leu Val Val Asn His Thr Ser Ser Glu His Pro Trp Phe
145   150   155   160
Gln Ala Ala Leu Lys Asp Lys Asn Ser Lys Tyr Arg Asp Tyr Tyr Ile
                165   170   175
Trp Ala Asp Lys Asn Thr Asp Leu Asn Glu Lys Gly Ser Trp Gly Gln
            180   185   190
Gln Val Trp His Lys Ala Pro Asn Gly Glu Tyr Phe Tyr Gly Thr Phe
        195   200   205
Trp Glu Gly Met Pro Asp Leu Asn Tyr Asp Asn Pro Glu Val Arg Lys
    210   215   220
Glu Met Ile Asn Val Gly Lys Phe Trp Leu Lys Gln Gly Val Asp Gly
225   230   235   240
Phe Arg Leu Asp Ala Ala Leu His Ile Phe Lys Gly Gln Thr Pro Glu
                245   250   255
Gly Ala Lys Lys Asn Ile Val Trp Trp Asn Glu Phe Arg Asp Ala Met
            260   265   270
Lys Lys Glu Asn Pro Asn Val Tyr Leu Thr Gly Glu Val Trp Asp Gln
        275   280   285
Pro Glu Val Val Ala Pro Tyr Tyr Gln Ser Leu Asp Ser Leu Phe Asn
    290   295   300
Phe Asp Leu Ala Gly Lys Ile Val Ser Ser Val Lys Ala Gly Asn Asp

305 310 315 320
Gln Gly Ile Ala Thr Ala Ala Ala Ala Thr Asp Glu Leu Phe Lys Ser
325 330 335
Tyr Asn Pro Asn Lys Ile Asp Gly Ile Phe Leu Thr Asn His Asp Gln
340 345 350
Asn Arg Val Met Ser Glu Leu Ser Gly Asp Val Asn Lys Ala Lys Ser
355 360 365
Ala Ala Ser Ile Leu Leu Thr Leu Pro Gly Asn Pro Tyr Ile Tyr Tyr
370 375 380
Gly Glu Glu Ile Gly Met Thr Gly Glu Lys Pro Asp Glu Leu Ile Arg
385 390 395 400
Glu Pro Phe Arg Trp Tyr Glu Gly Asn Gly Leu Gly Gln Thr Ser Trp
405 410 415
Glu Thr Pro Val Tyr Asn Lys Gly Gly Asn Gly Val Ser Val Glu Ala
420 425 430
Gln Thr Lys Gln Lys Asp Ser Leu Leu Asn His Tyr Arg Glu Met Ile
435 440 445
Arg Val Arg Gln Gln His Glu Glu Leu Val Lys Gly Thr Leu Gln Ser
450 455 460
Ile Ser Val Asp Ser Lys Glu Val Val Ala Tyr Ser Arg Thr Tyr Lys
465 470 475 480
Gly Asn Ser Ile Ser Val Tyr His Asn Ile Ser Asn Gln Pro Val Lys
485 490 495
Val Ser Val Ala Ala Lys Gly Lys Leu Ile Phe Ala Ser Glu Lys Gly
500 505 510
Ala Lys Lys Val Lys Asn Gln Leu Val Ile Pro Ala Asn Thr Thr Val
515 520 525
Leu Val Lys
530

<210> 191
<211> 1596
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 191
atgcaaacga ttgcaaaaaa aggggatgaa acgatgaaag ggaaaaaatg gacagcttta 60
gctctaacac tgccgctggc tgctagctta tcaacaggcg ttcacgccga aaccgtacat 120
aaaggtaaat ctccaacagc agataaaaac ggtgtctttt atgaagtgta tgtaaactct 180
ttttacgatg caaataaaga tggacatggt gacttaaaag gtcttacaca aaagttggac 240
tatttaaatg acggcaattc tcatacaaaa aatgatcttc aagtaaacgg gatttggatg 300
atgccagtca acccttctcc tagctatcat aaatatgatg taacggacta ttataacatt 360
gatccgcagt acggaaatct gcaagatttt cgcaagctga tgaaagaagc agacaaacga 420
gacgtaaaag tcattatgga ccttgttgtg aatcatacga gcagtgaaca cccttggttt 480
caagctgcgt taaaagataa aaacagcaag tacagagatt actatatttg ggctgataaa 540
aataccgact tgaatgaaaa aggatcttgg ggacaacaag tatggcataa agctccaaac 600
ggagagtatt tttacggaac gttctgggaa ggaatgcctg acttaaatta cgataaccct 660
gaagtaagaa aagaaatgat taacgtcgga aagttttggc taaagcaagg cgttgacggg 720
ttccgcttag atgctgcgct tcatattttt aaaggtcaaa cagctgaagg cgctaagaaa 780
aatatcctgt ggtggaatga gtttagagat gcgatgaaaa aagaaaatcc gaatgtatat 840
ctaacgggtg aagtatggga tcagcctgaa gtggtagctc cttattatca atcgcttgat 900
tctttattta attttgattt agcaggaaaa attgtcagct ctgtaaaagc aggaaatgat 960
caaggaatcg ccactgcagc agcagcaaca gatgaactgt tcaaatcata caatccaaac 1020
aaaattgatg gcatattctt aaccaaccat gaccaaaatc gcgtcatgag tgagctgagc 1080
ggcgatgtga gcaaagcaaa atcagctgct tctatcttac ttacgcttcc tggcaacccg 1140
tatatttatt acggtgaaga aatcggcatg accggtgaaa agcctgatga attaatccgt 1200
gaaccgttcc gctggtacga aggaaacgga cttggacaaa ccagttggga aacacctgta 1260
tacaataaag gcggaaacgg tgtgtctgta gaagcacaaa ccaaacaaaa ggattctttg 1320
ttaaatcatt accgtgaaat gattcgcgtg cgtcagcagc atgaagagtt agtaaaagga 1380
acgcttcaat ctatttcagt agacagtaaa gaagttgttg cttatagccg tacgtataaa 1440
ggcaactcca ttagtgtgta tcataatatt tcaaatcaac cggtaaaagt atctgtagca 1500
gcgaaaggta aattgatttt tgctagtgaa aaaggtgcta agaaagtcaa aaatcagctt 1560
gtggttccgg cgaatacaac ggttttaatg aaataa 1596

<210> 192

<211> 531
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 192

```
Met Gln Thr Ile Ala Lys Lys Gly Asp Glu Thr Met Lys Gly Lys Lys
 1               5                  10                  15
Trp Thr Ala Leu Ala Leu Thr Leu Pro Leu Ala Ala Ser Leu Ser Thr
            20                  25                  30
Gly Val His Ala Glu Thr Val His Lys Gly Lys Ser Pro Thr Ala Asp
        35                  40                  45
Lys Asn Gly Val Phe Tyr Glu Val Tyr Val Asn Ser Phe Tyr Asp Ala
    50                  55                  60
Asn Lys Asp Gly His Gly Asp Leu Lys Gly Leu Thr Gln Lys Leu Asp
65                  70                  75                  80
Tyr Leu Asn Asp Gly Asn Ser His Thr Lys Asn Asp Leu Gln Val Asn
                85                  90                  95
Gly Ile Trp Met Met Pro Val Asn Pro Ser Pro Ser Tyr His Lys Tyr
            100                 105                 110
Asp Val Thr Asp Tyr Tyr Asn Ile Asp Pro Gln Tyr Gly Asn Leu Gln
        115                 120                 125
Asp Phe Arg Lys Leu Met Lys Glu Ala Asp Lys Arg Asp Val Lys Val
    130                 135                 140
Ile Met Asp Leu Val Val Asn His Thr Ser Ser Glu His Pro Trp Phe
145                 150                 155                 160
Gln Ala Ala Leu Lys Asp Lys Asn Ser Lys Tyr Arg Asp Tyr Tyr Ile
                165                 170                 175
Trp Ala Asp Lys Asn Thr Asp Leu Asn Glu Lys Gly Ser Trp Gly Gln
            180                 185                 190
Gln Val Trp His Lys Ala Pro Asn Gly Glu Tyr Phe Tyr Gly Thr Phe
        195                 200                 205
Trp Glu Gly Met Pro Asp Leu Asn Tyr Asp Asn Pro Glu Val Arg Lys
    210                 215                 220
Glu Met Ile Asn Val Gly Lys Phe Trp Leu Lys Gln Gly Val Asp Gly
225                 230                 235                 240
Phe Arg Leu Asp Ala Ala Leu His Ile Phe Lys Gly Gln Thr Ala Glu
                245                 250                 255
Gly Ala Lys Lys Asn Ile Leu Trp Trp Asn Glu Phe Arg Asp Ala Met
            260                 265                 270
Lys Lys Glu Asn Pro Asn Val Tyr Leu Thr Gly Glu Val Trp Asp Gln
        275                 280                 285
Pro Glu Val Val Ala Pro Tyr Tyr Gln Ser Leu Asp Ser Leu Phe Asn
    290                 295                 300
Phe Asp Leu Ala Gly Lys Ile Val Ser Ser Val Lys Ala Gly Asn Asp
305                 310                 315                 320
Gln Gly Ile Ala Thr Ala Ala Ala Ala Thr Asp Glu Leu Phe Lys Ser
                325                 330                 335
Tyr Asn Pro Asn Lys Ile Asp Gly Ile Phe Leu Thr Asn His Asp Gln
            340                 345                 350
Asn Arg Val Met Ser Glu Leu Ser Gly Asp Val Ser Lys Ala Lys Ser
        355                 360                 365
Ala Ala Ser Ile Leu Leu Thr Leu Pro Gly Asn Pro Tyr Ile Tyr Tyr
    370                 375                 380
Gly Glu Glu Ile Gly Met Thr Gly Glu Lys Pro Asp Glu Leu Ile Arg
385                 390                 395                 400
Glu Pro Phe Arg Trp Tyr Glu Gly Asn Gly Leu Gly Gln Thr Ser Trp
                405                 410                 415
Glu Thr Pro Val Tyr Asn Lys Gly Gly Asn Gly Val Ser Val Glu Ala
            420                 425                 430
Gln Thr Lys Gln Lys Asp Ser Leu Leu Asn His Tyr Arg Glu Met Ile
        435                 440                 445
Arg Val Arg Gln Gln His Glu Glu Leu Val Lys Gly Thr Leu Gln Ser
    450                 455                 460
Ile Ser Val Asp Ser Lys Glu Val Val Ala Tyr Ser Arg Thr Tyr Lys
465                 470                 475                 480
```

Gly Asn Ser Ile Ser Val Tyr His Asn Ile Ser Asn Gln Pro Val Lys
485 490 495
Val Ser Val Ala Ala Lys Gly Lys Leu Ile Phe Ala Ser Glu Lys Gly
500 505 510
Ala Lys Lys Val Lys Asn Gln Leu Val Val Pro Ala Asn Thr Thr Val
515 520 525
Leu Met Lys
530

<210> 193
<211> 1962
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 193
atgaaattca aaaagagttt atctgccggg ctccttttgt tcggaggtct gagcggtgtg 60
acaccatccg tcgctgcgga ggtgccacga accgcatttg tccatttatt cgaatggagt 120
tggccggata ttgccaccga atgcgaaacc tttcttggcc ctaaggggtt ctctgcggtt 180
caggtgtctc cgccgcaaaa aagcgtcagc aatgctgcct ggtgggcgcg ctaccaacct 240
gttagttact cttttgaagg gcgcagtgga acccgggctc aatttgcgga tatggtccag 300
cgttgtaaag cggtgggggt cgatatttat ctggatgcgg tgatcaacca tatggcagca 360
caagatcgct attttccaga agtaccttac agcagtaatg attttcacag ttgcacgggc 420
gatatcgatt attccaaccg ctggtcgatt caaaattgcg atctggttgg gctgaacgat 480
ctcaaaaccg agtcagaata cgttcggcag aaaattgcag actatatgaa cgatgcgctc 540
agtctgggcg tggcggggtt tcggattgat gccgccaagc atatcccggc cggcgacatc 600
gcggcgatca agagcaagct caacggcagc ccgtatatct atcaggaggt tatcggggcg 660
gcaggggagc cggtacaaac cagcgagtac acgtatattg gagacgtgac ggaatttaac 720
ttcgcccgga ccatcgggcc taaatttaag caaggtaata ttaaagacct gcaggggatt 780
ggttcgtgga gcggctggct gagcagcgac gatgcggtga cctttgtgac caaccatgac 840
gaagaacgcc ataaccctgg ccaggttctc agccatcagg actttggcaa tctgtatttc 900
ctcggtaacg tgtttactct ggcgtatcct tacggctacc caaaagtgat gtcggggtac 960
tacttcagta attttgatgc cgggccacca tcgacagggg tacattctgg taatgcgtgt 1020
ggctttgatg gcggtgattg ggtctgcgaa cacaaatggc gtggtgtagc caacatggtg 1080
gcgtttcgca accacacagc agcccagtgg caggtcactg actggtggga cgatggttac 1140
aatcaggtgg cgtttggtcg tggcgggctg ggctttgtgg tgatcaatcg agatgacaat 1200
aaaggcatca atcagagttt ccagacggga atgcccgctg gcgagtattg tgacatcatt 1260
gccggtgatt tcgacaccca gagcggtcat tgcagcgcta cgacgatcac cgtcgacagt 1320
caggggtatg cacattttac tgtcggtagt catcaggccg ctgcgattca cattggcgcg 1380
aaactcggct ccgtgtgcca ggactgtggc ggcacggccg cagagacaaa agtctgcttt 1440
gacaatgcac aaaactttag ccaaccgtat ttgcattact ggaatgtcaa tgcggatcag 1500
gccgtagcga atgcaacctg gccgggcgtc gcgatgacgg ctgaaaatgg cggttactgc 1560
tacgattttg gtgtcggtct caattcactt caggtaattt tcagcgataa cggcgccagc 1620
caaaccgctg atctgaccgc cagcagtccg acgttgtgtt accagaacgg aacgtggcgt 1680
gacagtgact tctgtcagag tagcaatgtg ggcaacgaga gttggtattt ccgtggaacc 1740
tcaaacggtt ggggcgtgag cgcactcact tatgaggctg cgacaggcct gtacactacg 1800
gtgcagagct ttaacgggga ggagtcgccc gcacgcttta aaattgatga tggcaactgg 1860
agtgagtcgt atccaagtgc tgattatcaa gtcggtgatt atgccaccta cacgatcacg 1920
tttgacagcc agacgaaggc catcaccgtg acttcgcagt aa 1962

<210> 194
<211> 653
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 194
Met Lys Phe Lys Lys Ser Leu Ser Ala Gly Leu Leu Leu Phe Gly Gly
1 5 10 15
Leu Ser Gly Val Thr Pro Ser Val Ala Ala Glu Val Pro Arg Thr Ala
20 25 30
Phe Val His Leu Phe Glu Trp Ser Trp Pro Asp Ile Ala Thr Glu Cys
35 40 45
Glu Thr Phe Leu Gly Pro Lys Gly Phe Ser Ala Val Gln Val Ser Pro

```
    50                  55                  60
Pro Gln Lys Ser Val Ser Asn Ala Ala Trp Trp Ala Arg Tyr Gln Pro
65                  70                  75                  80
Val Ser Tyr Ser Phe Glu Gly Arg Ser Gly Thr Arg Ala Gln Phe Ala
                85                  90                  95
Asp Met Val Gln Arg Cys Lys Ala Val Gly Val Asp Ile Tyr Leu Asp
            100                 105                 110
Ala Val Ile Asn His Met Ala Ala Gln Asp Arg Tyr Phe Pro Glu Val
        115                 120                 125
Pro Tyr Ser Ser Asn Asp Phe His Ser Cys Thr Gly Asp Ile Asp Tyr
    130                 135                 140
Ser Asn Arg Trp Ser Ile Gln Asn Cys Asp Leu Val Gly Leu Asn Asp
145                 150                 155                 160
Leu Lys Thr Glu Ser Glu Tyr Val Arg Gln Lys Ile Ala Asp Tyr Met
                165                 170                 175
Asn Asp Ala Leu Ser Leu Gly Val Ala Gly Phe Arg Ile Asp Ala Ala
            180                 185                 190
Lys His Ile Pro Ala Gly Asp Ile Ala Ala Ile Lys Ser Lys Leu Asn
        195                 200                 205
Gly Ser Pro Tyr Ile Tyr Gln Glu Val Ile Gly Ala Ala Gly Glu Pro
    210                 215                 220
Val Gln Thr Ser Glu Tyr Thr Tyr Ile Gly Asp Val Thr Glu Phe Asn
225                 230                 235                 240
Phe Ala Arg Thr Ile Gly Pro Lys Phe Lys Gln Gly Asn Ile Lys Asp
                245                 250                 255
Leu Gln Gly Ile Gly Ser Trp Ser Gly Trp Leu Ser Ser Asp Asp Ala
            260                 265                 270
Val Thr Phe Val Thr Asn His Asp Glu Glu Arg His Asn Pro Gly Gln
        275                 280                 285
Val Leu Ser His Gln Asp Phe Gly Asn Leu Tyr Phe Leu Gly Asn Val
    290                 295                 300
Phe Thr Leu Ala Tyr Pro Tyr Gly Tyr Pro Lys Val Met Ser Gly Tyr
305                 310                 315                 320
Tyr Phe Ser Asn Phe Asp Ala Gly Pro Pro Ser Thr Gly Val His Ser
                325                 330                 335
Gly Asn Ala Cys Gly Phe Asp Gly Gly Asp Trp Val Cys Glu His Lys
            340                 345                 350
Trp Arg Gly Val Ala Asn Met Val Ala Phe Arg Asn His Thr Ala Ala
        355                 360                 365
Gln Trp Gln Val Thr Asp Trp Trp Asp Asp Gly Tyr Asn Gln Val Ala
    370                 375                 380
Phe Gly Arg Gly Gly Leu Gly Phe Val Val Ile Asn Arg Asp Asp Asn
385                 390                 395                 400
Lys Gly Ile Asn Gln Ser Phe Gln Thr Gly Met Pro Ala Gly Glu Tyr
                405                 410                 415
Cys Asp Ile Ile Ala Gly Asp Phe Asp Thr Gln Ser Gly His Cys Ser
            420                 425                 430
Ala Thr Thr Ile Thr Val Asp Ser Gln Gly Tyr Ala His Phe Thr Val
        435                 440                 445
Gly Ser His Gln Ala Ala Ala Ile His Ile Gly Ala Lys Leu Gly Ser
    450                 455                 460
Val Cys Gln Asp Cys Gly Gly Thr Ala Ala Glu Thr Lys Val Cys Phe
465                 470                 475                 480
Asp Asn Ala Gln Asn Phe Ser Gln Pro Tyr Leu His Tyr Trp Asn Val
                485                 490                 495
Asn Ala Asp Gln Ala Val Ala Asn Ala Thr Trp Pro Gly Val Ala Met
            500                 505                 510
Thr Ala Glu Asn Gly Gly Tyr Cys Tyr Asp Phe Gly Val Gly Leu Asn
        515                 520                 525
Ser Leu Gln Val Ile Phe Ser Asp Asn Gly Ala Ser Gln Thr Ala Asp
    530                 535                 540
Leu Thr Ala Ser Ser Pro Thr Leu Cys Tyr Gln Asn Gly Thr Trp Arg
545                 550                 555                 560
Asp Ser Asp Phe Cys Gln Ser Ser Asn Val Gly Asn Glu Ser Trp Tyr
                565                 570                 575
Phe Arg Gly Thr Ser Asn Gly Trp Gly Val Ser Ala Leu Thr Tyr Glu
            580                 585                 590
Ala Ala Thr Gly Leu Tyr Thr Thr Val Gln Ser Phe Asn Gly Glu Glu
```

595 600 605
Ser Pro Ala Arg Phe Lys Ile Asp Asp Gly Asn Trp Ser Glu Ser Tyr
610 615 620
Pro Ser Ala Asp Tyr Gln Val Gly Asp Tyr Ala Thr Tyr Thr Ile Thr
625 630 635 640
Phe Asp Ser Gln Thr Lys Ala Ile Thr Val Thr Ser Gln
645 650

<210> 195
<211> 2790
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 195
atgctgacag accgtttctt tgatggcgat acatcaaaca acgaccctta caaccagaac 60
tacgatgcta aaaacgaccg gggaacttat cagggcggcg attttaaagg aatcacgcaa 120
aaattggatt atctcgataa gctaggcgtg aacacaatct ggatcagccc gatcgtggaa 180
aatatcaagc atgatgtccg ttatgacaac tctgaagggc attcatacta tgcttaccac 240
ggctactggg caagcaactt cggtgcgtta aacccacact tcggtacaat ggaagatttc 300
catacactga ttgacgctgc ccatgaaaaa ggcatcaaga tcatggttga cgtagtatta 360
aaccacactg gttatggctt aaaagatatc aacggagaag tttccaatcc tccagccggt 420
tacccaactg acgcagaacg cagcacatat agcagcctgc ttcgccaggg ttcaaatgtc 480
ggctctgatg aggttgttgg cgaattagct ggcctacctg acttaaaaac agaagacccc 540
gcagtccgcc agacaatcat cgactggcaa acagactgga tcacgaaagc tactacagct 600
aaaggaaaca caattgacta cttccgtgtc gacactgtga agcacgttga agacgcaaca 660
tggatggcat tcaaaaatga cctcactgaa aaaatgccga cacacaaaat gatcggggaa 720
gcttggggag caagtgccaa taaccaactt ggataccttg aaacaggtat gatggactca 780
ctgcttgact tcgacttcaa aggcattgcg cacgatttcg tgaacggcaa gcttaaggca 840
gcaaacgatg ccctgactgc ccgcaacggt aaaattgaca acacagctac tttaggttca 900
ttccttggaa gccatgacga agatggtttc ctatttaaag aaggaaatga caaaggcaag 960
cttaaggttg ctgcttccct gcaagcaaca tcaaaaggcc agccggtcat ctattatggt 1020
gaagagcttg gtcaaagtgg agcaaacaac tatccgcaat acgataaccg ttatgacctg 1080
gcatgggaca aagttgaaaa caacgacgtc cttgagcact acactaaggt cctgaacttc 1140
agaagcgctc attcagaagt gttcgctaaa ggtgaacgcg caacaattgg cggttctgac 1200
gctgataaat tcttactttt tgctcgtaaa aatggaaacg aagctgctta cgtcggcttg 1260
aacgttgctg acacagcaaa agacgtaaca ctgactgttt ctgcaggtgc agtcgtaact 1320
gaccactatg cagataaaac ttatactgct tcagaagctg gagaaatcac attgacgatc 1380
ccggcaaaag ctgatggcgg tactgtttta ctaacggttg aaggcggaga aatcacagct 1440
gctaaagcgg caagcgaagg cgacggcaca gttgagccag tccctgcgaa ccacatccgc 1500
attcactaca accgtacaga caacaactat gaaaactacg gtgcatggct gtggaacgat 1560
gtagcctccc cttctgccaa ctggccgact ggcgctacaa tgtttgaaaa aacagacagc 1620
tacggtgcat acatcgacgt accacttaaa gagggcgcta agaacatcgg cttcctcgtt 1680
atggatgtaa caaaaggtga tcagggtaaa gacggcggcg acaaaggttt tacgatctca 1740
tcacctgaaa tgaacgaaat ttggatcaag caaggttctg acaaggtgta cacttacgag 1800
ccagttgatc ttccggcgaa cactgtccgc gtccactatg tacgtgacaa cgcagactac 1860
gaaaacttcg gtatctggaa ctggggcgat gtaacagcac cttccgaaaa ctggcctaca 1920
ggcgcagcga aattcgatgg tacagaccgt tacggtgcgt atgtcgacat tacgctaaaa 1980
gaaggcgcaa agaacattgg aatgattgct cttaacactg caaatggaga gaaagacggc 2040
ggagataaat ccttcaacct tctggataaa tataatcgca tttggattaa acaaggtgat 2100
gacaatgtct acgtttctcc atactgggag caggcaacag gaatcaccaa tgcagaggta 2160
atctctgaag atacgattct attaggcttc acaatgactg acggcttaac acctgaatct 2220
ttaaaaggag gtcttgtaat taaagattca actggtgctg aagttgccat cgaaagtgct 2280
gaaatcacaa gcgcaacctc tgtaaaagta aaagcaacat tcgatttaga aaagcttcca 2340
ttatccatca catacgcagg cagaacagtt tcagcttcaa ctggctggag aatgcttgat 2400
gaaatgtacg cttatgatgg aaacgacctt ggtgcgactt acaaggacgg agcagcgacg 2460
cttaaattat gggctccgaa agcgagcaag gtaaccgcta acttctttga taaaaataat 2520
gccgctgaaa aaatcggcag cgtcgagtta acgaagggtg aaaaaggagt ctggtcagct 2580
atggttgctc ctggcgacct gaacgtaacc gatcttgaag gttattttta ccagtatgat 2640
gtaacaaatg acggtataac tcgccaggtg ttagatcctt atgcaaaatc aatggcagcc 2700
tttactgtga atacagaagg caatgctggt cctgacgggg acactgttgg caaggcggca 2760
attcaaaaag cttctcgaga gtacttctag 2790

<210> 196
<211> 929
<212> PRT

<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 196

```
Met Leu Thr Asp Arg Phe Phe Asp Gly Asp Thr Ser Asn Asn Asp Pro
 1               5                  10                  15
Tyr Asn Gln Asn Tyr Asp Ala Lys Asn Asp Arg Gly Thr Tyr Gln Gly
            20                  25                  30
Gly Asp Phe Lys Gly Ile Thr Gln Lys Leu Asp Tyr Leu Asp Lys Leu
        35                  40                  45
Gly Val Asn Thr Ile Trp Ile Ser Pro Ile Val Glu Asn Ile Lys His
    50                  55                  60
Asp Val Arg Tyr Asp Asn Ser Glu Gly His Ser Tyr Tyr Ala Tyr His
65                  70                  75                  80
Gly Tyr Trp Ala Ser Asn Phe Gly Ala Leu Asn Pro His Phe Gly Thr
                85                  90                  95
Met Glu Asp Phe His Thr Leu Ile Asp Ala Ala His Glu Lys Gly Ile
            100                 105                 110
Lys Ile Met Val Asp Val Val Leu Asn His Thr Gly Tyr Gly Leu Lys
        115                 120                 125
Asp Ile Asn Gly Glu Val Ser Asn Pro Pro Ala Gly Tyr Pro Thr Asp
    130                 135                 140
Ala Glu Arg Ser Thr Tyr Ser Ser Leu Leu Arg Gln Gly Ser Asn Val
145                 150                 155                 160
Gly Ser Asp Glu Val Val Gly Glu Leu Ala Gly Leu Pro Asp Leu Lys
                165                 170                 175
Thr Glu Asp Pro Ala Val Arg Gln Thr Ile Ile Asp Trp Gln Thr Asp
            180                 185                 190
Trp Ile Thr Lys Ala Thr Thr Ala Lys Gly Asn Thr Ile Asp Tyr Phe
        195                 200                 205
Arg Val Asp Thr Val Lys His Val Glu Asp Ala Thr Trp Met Ala Phe
    210                 215                 220
Lys Asn Asp Leu Thr Glu Lys Met Pro Thr His Lys Met Ile Gly Glu
225                 230                 235                 240
Ala Trp Gly Ala Ser Ala Asn Asn Gln Leu Gly Tyr Leu Glu Thr Gly
                245                 250                 255
Met Met Asp Ser Leu Leu Asp Phe Asp Phe Lys Gly Ile Ala His Asp
            260                 265                 270
Phe Val Asn Gly Lys Leu Lys Ala Ala Asn Asp Ala Leu Thr Ala Arg
        275                 280                 285
Asn Gly Lys Ile Asp Asn Thr Ala Thr Leu Gly Ser Phe Leu Gly Ser
    290                 295                 300
His Asp Glu Asp Gly Phe Leu Phe Lys Glu Gly Asn Asp Lys Gly Lys
305                 310                 315                 320
Leu Lys Val Ala Ala Ser Leu Gln Ala Thr Ser Lys Gly Gln Pro Val
                325                 330                 335
Ile Tyr Tyr Gly Glu Glu Leu Gly Gln Ser Gly Ala Asn Asn Tyr Pro
            340                 345                 350
Gln Tyr Asp Asn Arg Tyr Asp Leu Ala Trp Asp Lys Val Glu Asn Asn
        355                 360                 365
Asp Val Leu Glu His Tyr Thr Lys Val Leu Asn Phe Arg Ser Ala His
    370                 375                 380
Ser Glu Val Phe Ala Lys Gly Glu Arg Ala Thr Ile Gly Gly Ser Asp
385                 390                 395                 400
Ala Asp Lys Phe Leu Leu Phe Ala Arg Lys Asn Gly Asn Glu Ala Ala
                405                 410                 415
Tyr Val Gly Leu Asn Val Ala Asp Thr Ala Lys Asp Val Thr Leu Thr
            420                 425                 430
Val Ser Ala Gly Ala Val Val Thr Asp His Tyr Ala Asp Lys Thr Tyr
        435                 440                 445
Thr Ala Ser Glu Ala Gly Glu Ile Thr Leu Thr Ile Pro Ala Lys Ala
    450                 455                 460
Asp Gly Gly Thr Val Leu Leu Thr Val Glu Gly Gly Glu Ile Thr Ala
465                 470                 475                 480
Ala Lys Ala Ala Ser Glu Gly Asp Gly Thr Val Glu Pro Val Pro Ala
                485                 490                 495
```

Asn His Ile Arg Ile His Tyr Asn Arg Thr Asp Asn Asn Tyr Glu Asn
500 505 510
Tyr Gly Ala Trp Leu Trp Asn Asp Val Ala Ser Pro Ser Ala Asn Trp
515 520 525
Pro Thr Gly Ala Thr Met Phe Glu Lys Thr Asp Ser Tyr Gly Ala Tyr
530 535 540
Ile Asp Val Pro Leu Lys Glu Gly Ala Lys Asn Ile Gly Phe Leu Val
545 550 555 560
Met Asp Val Thr Lys Gly Asp Gln Gly Lys Asp Gly Gly Asp Lys Gly
565 570 575
Phe Thr Ile Ser Ser Pro Glu Met Asn Glu Ile Trp Ile Lys Gln Gly
580 585 590
Ser Asp Lys Val Tyr Thr Tyr Glu Pro Val Asp Leu Pro Ala Asn Thr
595 600 605
Val Arg Val His Tyr Val Arg Asp Asn Ala Asp Tyr Glu Asn Phe Gly
610 615 620
Ile Trp Asn Trp Gly Asp Val Thr Ala Pro Ser Glu Asn Trp Pro Thr
625 630 635 640
Gly Ala Ala Lys Phe Asp Gly Thr Asp Arg Tyr Gly Ala Tyr Val Asp
645 650 655
Ile Thr Leu Lys Glu Gly Ala Lys Asn Ile Gly Met Ile Ala Leu Asn
660 665 670
Thr Ala Asn Gly Glu Lys Asp Gly Gly Asp Lys Ser Phe Asn Leu Leu
675 680 685
Asp Lys Tyr Asn Arg Ile Trp Ile Lys Gln Gly Asp Asp Asn Val Tyr
690 695 700
Val Ser Pro Tyr Trp Glu Gln Ala Thr Gly Ile Thr Asn Ala Glu Val
705 710 715 720
Ile Ser Glu Asp Thr Ile Leu Leu Gly Phe Thr Met Thr Asp Gly Leu
725 730 735
Thr Pro Glu Ser Leu Lys Gly Gly Leu Val Ile Lys Asp Ser Thr Gly
740 745 750
Ala Glu Val Ala Ile Glu Ser Ala Glu Ile Thr Ser Ala Thr Ser Val
755 760 765
Lys Val Lys Ala Thr Phe Asp Leu Glu Lys Leu Pro Leu Ser Ile Thr
770 775 780
Tyr Ala Gly Arg Thr Val Ser Ala Ser Thr Gly Trp Arg Met Leu Asp
785 790 795 800
Glu Met Tyr Ala Tyr Asp Gly Asn Asp Leu Gly Ala Thr Tyr Lys Asp
805 810 815
Gly Ala Ala Thr Leu Lys Leu Trp Ala Pro Lys Ala Ser Lys Val Thr
820 825 830
Ala Asn Phe Phe Asp Lys Asn Asn Ala Ala Glu Lys Ile Gly Ser Val
835 840 845
Glu Leu Thr Lys Gly Glu Lys Gly Val Trp Ser Ala Met Val Ala Pro
850 855 860
Gly Asp Leu Asn Val Thr Asp Leu Glu Gly Tyr Phe Tyr Gln Tyr Asp
865 870 875 880
Val Thr Asn Asp Gly Ile Thr Arg Gln Val Leu Asp Pro Tyr Ala Lys
885 890 895
Ser Met Ala Ala Phe Thr Val Asn Thr Glu Gly Asn Ala Gly Pro Asp
900 905 910
Gly Asp Thr Val Gly Lys Ala Ala Ile Gln Lys Ala Ser Arg Glu Tyr
915 920 925
Phe

<210> 197
<211> 1401
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 197
atgaaaccgt caaaattcgt ttttctctct gctgccatcg cttgcagcct ctccagtacc 60
gccaatgctg acgccatttt gcatgcattt aactggaagt actccgacgt cacgcaaaac 120

```
gcctcgcaaa tcgcggcggc gggttataaa aaagtgctga tttcgccagc actgaaatcg     180
agtggcaatg aatggtgggc acgttatcaa ccgcaagatc tgcgcgtgat cgattcccca     240
cttggcaaca aaagtgactt aaaatccatg attgatgctc tgaaggcggt cggcgttgat     300
gtgtatgccg atgtggtgct taaccatatg gccaatgaaa catggaagcg tgaagactta     360
aattaccctg gcagtgaagt gctgcaacaa tacgcagcta acaccagtta ttatgcggac     420
caaacgcttt ttggcaattt aacggaaaac ctattctctg gctttgactt ccacccagaa     480
ggctgtatta gcgattggaa tgatgccggc aatgttcagt actggcgtct ttgtggcggt     540
gctggtgacc gagggctgcc agacttagat ccgaacaact gggtggtgtc acagcaacgt     600
ttgtatttga atgcgctaaa aggtttaggt gtgaaaggct tccgcattga tgcggttaaa     660
cacatgagcc aatatcaaat cgaccagatt ttcactgcag agattaccgc cggaatgcac     720
gtgtttggtg aagtgatcac cagtggtggc aaaggcgact ccagctatga gaacttctta     780
gcgccttatc tcaacgccac caaccattcg gcttacgatt tcccactgtt tgcctctatt     840
cgcaacgcct tctcctacag cggtggcatg aacatgcttc atgatccaca agcctatggc     900
caagggcttg aaaacgcacg ttcaattacc tttaccatca cgcacgacat cccaacgaac     960
gacggtttcc gttatcaaat catggatccg aaagatgaag agctggctta cgcttatatc    1020
ctcggtaaag atggcggcac acctctgatt tacagcgaca acttacctga taacgaagat    1080
cgtgataatc gccgttggga aggtgtttgg aaccgtgacc tgatgaagaa catgttgcgc    1140
ttccataacc aaatgcaagg gcaagagatg acgatgctgt acagcgacca atgtctactg    1200
atgtttaagc gcggtaaaca aggggtggtc ggcattaata aatgcggtga agagcgttct    1260
cataccgttg acacctatca gcatgagttc aactggtatc agccttacac agatacactc    1320
actggcgtga ctgaaaccgt gagttcgcgt taccacacct tccgaattcc agctcgcagc    1380
gcgcgcatgt acatgctcta a                                              1401

<210> 198
<211> 466
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 198
Met Lys Pro Ser Lys Phe Val Phe Leu Ser Ala Ala Ile Ala Cys Ser
 1               5                  10                  15
Leu Ser Ser Thr Ala Asn Ala Asp Ala Ile Leu His Ala Phe Asn Trp
            20                  25                  30
Lys Tyr Ser Asp Val Thr Gln Asn Ala Ser Gln Ile Ala Ala Ala Gly
        35                  40                  45
Tyr Lys Lys Val Leu Ile Ser Pro Ala Leu Lys Ser Ser Gly Asn Glu
    50                  55                  60
Trp Trp Ala Arg Tyr Gln Pro Gln Asp Leu Arg Val Ile Asp Ser Pro
65                  70                  75                  80
Leu Gly Asn Lys Ser Asp Leu Lys Ser Met Ile Asp Ala Leu Lys Ala
                85                  90                  95
Val Gly Val Asp Val Tyr Ala Asp Val Val Leu Asn His Met Ala Asn
            100                 105                 110
Glu Thr Trp Lys Arg Glu Asp Leu Asn Tyr Pro Gly Ser Glu Val Leu
        115                 120                 125
Gln Gln Tyr Ala Ala Asn Thr Ser Tyr Tyr Ala Asp Gln Thr Leu Phe
    130                 135                 140
Gly Asn Leu Thr Glu Asn Leu Phe Ser Gly Phe Asp Phe His Pro Glu
145                 150                 155                 160
Gly Cys Ile Ser Asp Trp Asn Asp Ala Gly Asn Val Gln Tyr Trp Arg
                165                 170                 175
Leu Cys Gly Gly Ala Gly Asp Arg Gly Leu Pro Asp Leu Asp Pro Asn
            180                 185                 190
Asn Trp Val Val Ser Gln Gln Arg Leu Tyr Leu Asn Ala Leu Lys Gly
        195                 200                 205
Leu Gly Val Lys Gly Phe Arg Ile Asp Ala Val Lys His Met Ser Gln
    210                 215                 220
Tyr Gln Ile Asp Gln Ile Phe Thr Ala Glu Ile Thr Ala Gly Met His
225                 230                 235                 240
Val Phe Gly Glu Val Ile Thr Ser Gly Gly Lys Gly Asp Ser Ser Tyr
                245                 250                 255
Glu Asn Phe Leu Ala Pro Tyr Leu Asn Ala Thr Asn His Ser Ala Tyr
            260                 265                 270
Asp Phe Pro Leu Phe Ala Ser Ile Arg Asn Ala Phe Ser Tyr Ser Gly
        275                 280                 285
```

```
Gly Met Asn Met Leu His Asp Pro Gln Ala Tyr Gly Gln Gly Leu Glu
    290                 295                 300
Asn Ala Arg Ser Ile Thr Phe Thr Ile Thr His Asp Ile Pro Thr Asn
305                 310                 315                 320
Asp Gly Phe Arg Tyr Gln Ile Met Asp Pro Lys Asp Glu Glu Leu Ala
                325                 330                 335
Tyr Ala Tyr Ile Leu Gly Lys Asp Gly Gly Thr Pro Leu Ile Tyr Ser
            340                 345                 350
Asp Asn Leu Pro Asp Asn Glu Asp Arg Asp Asn Arg Arg Trp Glu Gly
        355                 360                 365
Val Trp Asn Arg Asp Leu Met Lys Asn Met Leu Arg Phe His Asn Gln
    370                 375                 380
Met Gln Gly Gln Glu Met Thr Met Leu Tyr Ser Asp Gln Cys Leu Leu
385                 390                 395                 400
Met Phe Lys Arg Gly Lys Gln Gly Val Val Gly Ile Asn Lys Cys Gly
                405                 410                 415
Glu Glu Arg Ser His Thr Val Asp Thr Tyr Gln His Glu Phe Asn Trp
            420                 425                 430
Tyr Gln Pro Tyr Thr Asp Thr Leu Thr Gly Val Thr Glu Thr Val Ser
        435                 440                 445
Ser Arg Tyr His Thr Phe Arg Ile Pro Ala Arg Ser Ala Arg Met Tyr
    450                 455                 460
Met Leu
465
```

<210> 199
<211> 399
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 199

```
gtgagtttga ccaaaaaggc tcagtacgaa ccaaatacgg caccaaggct cagtacatct      60
ctgcaatcaa tgccgcgcac aacaacaata tccaaattta cggcgatgtt gtgtttaacc     120
accgaggtgg tgctgatggg aagtcgtggg tcgataccaa gcgcgttgat tgggacaacc     180
gcaatattga actgggcgac aaatggattg aagcttgggt tgagtttaat tttcctggcc     240
gcaacgacaa atactcgaac ttccattgga cttggtatca ctttgacggt gttgactggg     300
atgacgccgg caaagaaaaa gcgatcttta aattcaaagg cgaaggaaaa gcatgggatt     360
gggaagtcag ctctgaaaaa ggcaattacg actacctaa                            399
```

<210> 200
<211> 132
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 200

```
Val Ser Leu Thr Lys Lys Ala Gln Tyr Glu Pro Asn Thr Ala Pro Arg
 1                  5                  10                  15
Leu Ser Thr Ser Leu Gln Ser Met Pro Arg Thr Thr Thr Ile Ser Lys
            20                  25                  30
Phe Thr Ala Met Leu Cys Leu Thr Thr Glu Val Val Leu Met Gly Ser
        35                  40                  45
Arg Gly Ser Ile Pro Ser Ala Leu Ile Gly Thr Thr Ala Ile Leu Asn
    50                  55                  60
Trp Ala Thr Asn Gly Leu Lys Leu Gly Leu Ser Leu Ile Phe Leu Ala
65                  70                  75                  80
Ala Thr Thr Asn Thr Arg Thr Ser Ile Gly Leu Gly Ile Thr Leu Thr
                85                  90                  95
Val Leu Thr Gly Met Thr Pro Ala Lys Lys Lys Arg Ser Leu Asn Ser
            100                 105                 110
Lys Ala Lys Glu Lys His Gly Ile Gly Lys Ser Ala Leu Lys Lys Ala
        115                 120                 125
Ile Thr Thr Thr
```

130

<210> 201
<211> 1911
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 201
atgacagcca aggctgatga cttacgcatt taccagatca tggtggaaag ctttgtggat 60
ggcgataaac aggtcggcca tggcaccggc tacggtacca gccatcacaa aggcgatctg 120
caagggatca ttgactcgct ggattacatt caatcgctgg gcgtcaatgc catttggcta 180
acgccgattt ttgaatctat tccggtggag ggacaagacc attgggcgga caggcttgat 240
gctacaggct actttgccag tgactatttc aagatagacc ctcgctttgg cacgttagaa 300
caagcccgtg agctggtgga aaaggcacac gcgaaaggct tgtatgtctt ctttgatgga 360
gtatttggtc accataaagg caatgtggtg ccatcaccac aaggtagact gcctgtcggt 420
gaaaataacc cggtcagcta cccagagagc ctggcgtttt acgaagaagt cgccagttac 480
tgggtgaaag agttaaagat tgatggctgg cgtctggatc aagcctatca agtgccgacc 540
gatgcatgga aagcgatccg tcagagcgtt gatgaagcgt cacagtccgt aacttatgtg 600
aataacaaag gggaaaccgt ccatcctttg ggttacatgg tggctgaaat ttggaataac 660
gaacgttaca tcacagaaac cggttacggc aaagaaggcg atccggcgtt gtgctcggct 720
tttgattttc cgatgcgttt ccgagtggtc gaaacctttg cggttaacga aagtggtgtc 780
agccgaaaag gcggcgaatg gttgaatgac ggcatgtcac tgcacagtca gtatccggat 840
catgccaagc ctaatttaat gttgggcaac catgatgtgg tgcgctttgg ggatctgctg 900
caacgtggcg gtattgcgtc accagaacaa ccgcaatact ggcagcgtca taaagcggcg 960
atgtctttct tagcagcgta taccggccca ttaccttgt attacggtga agaaattggc 1020
gatcaggttg acggctttgc taaaaaaatc aaagaagatt gtgccgttat tggtttgtgt 1080
gatgaccacg tggcgcgcac cagtgcgaag attgatggcg tgacggcgtc actgaatgca 1140
cagcagtctg aactcaaagt atatgtctct tcattgatga cattacgtca gcaacatcct 1200
gcgttatcac aaggggaacg tactaatgtg atggcgacag agacagtata cgtagaccat 1260
aaacaggcag acaatgaagc cctgttgtac atggtgagta cgactgataa cgcggagtca 1320
gtcaccttga agggcaaagc gattggttca caaggtgtgc tgattgattt gttaacgaac 1380
gagcgtttta tgcccaataa tggggagtat gccattccat taacgggctt tggcgcacga 1440
ttcctcaaga ttgacactcc gacagcggcg ggtgtgatgg cgcaatctgc tgcctcggta 1500
tcgctagtag gtgaagggat catggcccaa tgtgataccc caaccgttga aggcaccggt 1560
ccggtagcag aaaccttgta cgtggttggc gattttgccg atgctggttg gaagcaaaag 1620
ccgcagcgcg cgtatcaata caaaggcaag cacaatggca gcaacttgta tcaagtggtt 1680
gtcgatgaaa aagcgggcgc ctacaagatg caatacgcca cgaaagattg gagcccacag 1740
tttactgcag acggtatggc attgaagccg ggtaccgcaa agtcgctcat agcgggtggc 1800
tacggtaaag acaccgccgt gacgttgccg gaatccggta agtatgtgtg gagcttaaca 1860
ttcagtgatc ttggcgagcc ggagcaaatc atggtgtcta agtgtcagta a 1911

<210> 202
<211> 630
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 202
Met Thr Ala Lys Ala Asp Asp Leu Arg Ile Tyr Gln Ile Met Val Glu
1 5 10 15
Ser Phe Val Asp Gly Asp Lys Gln Val Gly His Gly Thr Gly Tyr Gly
20 25 30
Thr Ser His His Lys Gly Asp Leu Gln Gly Ile Ile Asp Ser Leu Asp
35 40 45
Tyr Ile Gln Ser Leu Gly Val Asn Ala Ile Trp Leu Thr Pro Ile Phe
50 55 60
Glu Ser Ile Pro Val Glu Gly Gln Asp His Trp Ala Asp Arg Leu Asp
65 70 75 80
Ala Thr Gly Tyr Phe Ala Ser Asp Tyr Phe Lys Ile Asp Pro Arg Phe
85 90 95
Gly Thr Leu Glu Gln Ala Arg Glu Leu Val Glu Lys Ala His Ala Lys
100 105 110
Gly Leu Tyr Val Phe Phe Asp Gly Val Phe Gly His His Lys Gly Asn

```
        115                 120                 125
Val Val Pro Ser Pro Gln Gly Arg Leu Pro Val Gly Glu Asn Asn Pro
    130                 135                 140
Val Ser Tyr Pro Glu Ser Leu Ala Phe Tyr Glu Glu Val Ala Ser Tyr
145                 150                 155                 160
Trp Val Lys Glu Leu Lys Ile Asp Gly Trp Arg Leu Asp Gln Ala Tyr
                165                 170                 175
Gln Val Pro Thr Asp Ala Trp Lys Ala Ile Arg Gln Ser Val Asp Glu
            180                 185                 190
Ala Ser Gln Ser Val Thr Tyr Val Asn Asn Lys Gly Glu Thr Val His
        195                 200                 205
Pro Leu Gly Tyr Met Val Ala Glu Ile Trp Asn Asn Glu Arg Tyr Ile
    210                 215                 220
Thr Glu Thr Gly Tyr Gly Lys Glu Gly Asp Pro Ala Leu Cys Ser Ala
225                 230                 235                 240
Phe Asp Phe Pro Met Arg Phe Arg Val Val Glu Thr Phe Ala Val Asn
                245                 250                 255
Glu Ser Gly Val Ser Arg Lys Gly Gly Glu Trp Leu Asn Asp Gly Met
            260                 265                 270
Ser Leu His Ser Gln Tyr Pro Asp His Ala Lys Pro Asn Leu Met Leu
        275                 280                 285
Gly Asn His Asp Val Val Arg Phe Gly Asp Leu Leu Gln Arg Gly Gly
    290                 295                 300
Ile Ala Ser Pro Glu Gln Pro Gln Tyr Trp Gln Arg His Lys Ala Ala
305                 310                 315                 320
Met Ser Phe Leu Ala Ala Tyr Thr Gly Pro Ile Thr Leu Tyr Tyr Gly
                325                 330                 335
Glu Glu Ile Gly Asp Gln Val Asp Gly Phe Ala Lys Lys Ile Lys Glu
            340                 345                 350
Asp Cys Ala Val Ile Gly Leu Cys Asp Asp His Val Ala Arg Thr Ser
        355                 360                 365
Ala Lys Ile Asp Gly Val Thr Ala Ser Leu Asn Ala Gln Gln Ser Glu
    370                 375                 380
Leu Lys Val Tyr Val Ser Ser Leu Met Thr Leu Arg Gln Gln His Pro
385                 390                 395                 400
Ala Leu Ser Gln Gly Glu Arg Thr Asn Val Met Ala Thr Glu Thr Val
                405                 410                 415
Tyr Val Asp His Lys Gln Ala Asp Asn Glu Ala Leu Leu Tyr Met Val
            420                 425                 430
Ser Thr Thr Asp Asn Ala Glu Ser Val Thr Leu Lys Gly Lys Ala Ile
        435                 440                 445
Gly Ser Gln Gly Val Leu Ile Asp Leu Leu Thr Asn Glu Arg Phe Met
    450                 455                 460
Pro Asn Asn Gly Glu Tyr Ala Ile Pro Leu Thr Gly Phe Gly Ala Arg
465                 470                 475                 480
Phe Leu Lys Ile Asp Thr Pro Thr Ala Ala Gly Val Met Ala Gln Ser
                485                 490                 495
Ala Ala Ser Val Ser Leu Val Gly Glu Gly Ile Met Ala Gln Cys Asp
            500                 505                 510
Thr Pro Thr Val Glu Gly Thr Gly Pro Val Ala Glu Thr Leu Tyr Val
        515                 520                 525
Val Gly Asp Phe Ala Asp Ala Gly Trp Lys Gln Lys Pro Gln Arg Ala
    530                 535                 540
Tyr Gln Tyr Lys Gly Lys His Asn Gly Ser Asn Leu Tyr Gln Val Val
545                 550                 555                 560
Val Asp Glu Lys Ala Gly Ala Tyr Lys Met Gln Tyr Ala Thr Lys Asp
                565                 570                 575
Trp Ser Pro Gln Phe Thr Ala Asp Gly Met Ala Leu Lys Pro Gly Thr
            580                 585                 590
Ala Lys Ser Leu Ile Ala Gly Gly Tyr Gly Lys Asp Thr Ala Val Thr
        595                 600                 605
Leu Pro Glu Ser Gly Lys Tyr Val Trp Ser Leu Thr Phe Ser Asp Leu
    610                 615                 620
Gly Glu Pro Glu Gln Ile
625                 630

<210> 203
<211> 2601
```

<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 203
atgaagatga agtcccgggc gtggttgtta ggtagtgcag tggccatggc gttggcctct 60
tcggcagcca atgccggtgt catggttcac ctgttccagt ggaagtacaa tgacatcgcc 120
aacgagtgcg aaaaggtgct cggtcccaaa gggtatgaag cagtgcagat cacgccgcct 180
gctgaacacc tgcaaggctc ctcctggtgg gtggtctatc agcccgtcag ctacaagaac 240
ttcacttctc tgggcggtaa cgaggccgaa ctcaaaagca tgatcgcccg ttgcaaggcc 300
gccggggtca agatttacgc cgatgcggta ttcaaccagc tggctggtgg atcaggcgtc 360
ggtacaggtg gtagcagcta caatgccggc agcttcagct atccccaatt tggctacaac 420
gatttccatc acgctgggag cctcaccaac tatgccgacc gcaacaatgt gcaaaacggt 480
gccctgctgg ggctgccgga tctggatacc ggctctgcct atgtgcagga tcagctggct 540
acctatatga agaccctgag tggctggggt gtggcaggtt ttcgtcttga tgcagcaaag 600
catatgagcg ttgccgatct ctcggccatc gtcagcaagg cgggcaatcc ttttgtctac 660
tccgaggtga ttggtgccac gggtgaacca atccagccgg gcgaatatac cggcattggt 720
gccgtgaccg aatttaaata cggcaccgat ctggcctcca acttcaaggg gcagatcaag 780
aatctcaaga gcatgggcga gagctggggt ctgcttgcgt cgaacaaggc tgaagtcttt 840
gtggtcaacc atgaccgtga gcggggacat ggcggtggcg gtatgctgac ctacaaggat 900
ggtgccctct acaatctggc caacatcttc atgctggcct ggccctatgg cgcctatccc 960
caggtgatgt ccggctatga tttcggcacc aataccgata ttggtgggcc gagcgctacc 1020
ccttgttctt ccggctctag ctggaactgc gaacaccgct ggagcaacat cgccaacatg 1080
gtctcgttcc acaatgccgc ccaaggcacg tccatgacca actggtggga taatggtaat 1140
aaccagatcg cctttggtcg cggcgccaag gcctttgtgg tgatcaacaa tgaatcttcc 1200
actctgagca agagcctgca gacgggtctg ccagccgggg agtactgcaa cattctggcc 1260
ggtgatgccc tgtgcagcgg cagcaccatc aaggtggatg ccagcggtat ggccaccttc 1320
aacgtggcag ggatgaaggc ggcagcgatc catatcaatg ccaagcccga tagcaccagc 1380
agtggcagct caggctcttc ctctggctct tcttcctctg ccaccagtaa caagtttgcc 1440
agcatgaatc tgcggggcac caacaatggc tgggccagca ccgccatgac agtggatgcc 1500
aaccgtgtct ggtcggcgga tgtcaccttt accggggccg cggatgccaa tggtgcccag 1560
cgcttcaagt ttgatgtcta tggcaactgg acagagagct atggcgatac acaagccgat 1620
ggcattgccg acaaggggag cgccaaggac atctatttca atggtgtggg caagtatcgt 1680
gtctcgctca aggagagcga catgagctac accctgaccc agctctccag caatcaggca 1740
ccggtggcgg ccatcacccc caagacactc tccgtcaagc tgggtgactc agtggtgttc 1800
gatgcctccg gctccaccga tgatgtgggt gtcactggct acagctggtc taccggtggc 1860
agtgccaaga ccgaaactgt gctgtttgat gctctgggta ccaagaccat taccgtgaca 1920
gtggccgatg ccgatggctt gacctccaag gccagtgcca ccgtcaccgt caccgatggc 1980
agcgtggctt ataacagcaa ctttgccagc ctgaacttcc gtggcactcc caacagttgg 2040
ggcgcggcag ccatgacgct ggtggcagac aacacctggg aggcaacggt caacttcgat 2100
ggtcaggcca atcagcgctt caagttcgat atcaagggtg actggagcca gaactatggt 2160
gatagcaaca aggatggggt ggccgaacgt accggtgccg atatttacac cactgtgacc 2220
ggtcaatata aggtgcaatt taacgactcc actttgaagt acaccctgac caagctggcc 2280
gatagcagcg ccaccagcta tagcgcgaac tttgccagcc tctacctgcg tggcaccccg 2340
aacagctggg gcaccaccgc catgaagctg gtggccaata acagctggca ggccgaggtg 2400
accttcaccg gcaagggcga tgccactggt gcccaacgct tcaagttcga cgtcaagggt 2460
gactggagcc agaactacgg tgacagcaac atggacggga ctgccgaacg gactggtggc 2520
gatatcacca gtgccgtggt gggcacctat ctggtgacct ttaatgacag cacactgaaa 2580
tacaccctga ccgccaaata a 2601

<210> 204
<211> 866
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 204
Met Lys Met Lys Ser Arg Ala Trp Leu Leu Gly Ser Ala Val Ala Met
1 5 10 15
Ala Leu Ala Ser Ser Ala Ala Asn Ala Gly Val Met Val His Leu Phe
20 25 30
Gln Trp Lys Tyr Asn Asp Ile Ala Asn Glu Cys Glu Lys Val Leu Gly
35 40 45
Pro Lys Gly Tyr Glu Ala Val Gln Ile Thr Pro Pro Ala Glu His Leu

```
    50                  55                  60
Gln Gly Ser Ser Trp Trp Val Val Tyr Gln Pro Val Ser Tyr Lys Asn
65                  70                  75                  80
Phe Thr Ser Leu Gly Gly Asn Glu Ala Glu Leu Lys Ser Met Ile Ala
                85                  90                  95
Arg Cys Lys Ala Ala Gly Val Lys Ile Tyr Ala Asp Ala Val Phe Asn
            100                 105                 110
Gln Leu Ala Gly Gly Ser Gly Val Gly Thr Gly Gly Ser Ser Tyr Asn
        115                 120                 125
Ala Gly Ser Phe Ser Tyr Pro Gln Phe Gly Tyr Asn Asp Phe His His
    130                 135                 140
Ala Gly Ser Leu Thr Asn Tyr Ala Asp Arg Asn Asn Val Gln Asn Gly
145                 150                 155                 160
Ala Leu Leu Gly Leu Pro Asp Leu Asp Thr Gly Ser Ala Tyr Val Gln
                165                 170                 175
Asp Gln Leu Ala Thr Tyr Met Lys Thr Leu Ser Gly Trp Gly Val Ala
            180                 185                 190
Gly Phe Arg Leu Asp Ala Ala Lys His Met Ser Val Ala Asp Leu Ser
        195                 200                 205
Ala Ile Val Ser Lys Ala Gly Asn Pro Phe Val Tyr Ser Glu Val Ile
    210                 215                 220
Gly Ala Thr Gly Glu Pro Ile Gln Pro Gly Glu Tyr Thr Gly Ile Gly
225                 230                 235                 240
Ala Val Thr Glu Phe Lys Tyr Gly Thr Asp Leu Ala Ser Asn Phe Lys
                245                 250                 255
Gly Gln Ile Lys Asn Leu Lys Ser Met Gly Glu Ser Trp Gly Leu Leu
            260                 265                 270
Ala Ser Asn Lys Ala Glu Val Phe Val Val Asn His Asp Arg Glu Arg
        275                 280                 285
Gly His Gly Gly Gly Gly Met Leu Thr Tyr Lys Asp Gly Ala Leu Tyr
    290                 295                 300
Asn Leu Ala Asn Ile Phe Met Leu Ala Trp Pro Tyr Gly Ala Tyr Pro
305                 310                 315                 320
Gln Val Met Ser Gly Tyr Asp Phe Gly Thr Asn Thr Asp Ile Gly Gly
                325                 330                 335
Pro Ser Ala Thr Pro Cys Ser Ser Gly Ser Ser Trp Asn Cys Glu His
            340                 345                 350
Arg Trp Ser Asn Ile Ala Asn Met Val Ser Phe His Asn Ala Ala Gln
        355                 360                 365
Gly Thr Ser Met Thr Asn Trp Trp Asp Asn Gly Asn Asn Gln Ile Ala
    370                 375                 380
Phe Gly Arg Gly Ala Lys Ala Phe Val Val Ile Asn Asn Glu Ser Ser
385                 390                 395                 400
Thr Leu Ser Lys Ser Leu Gln Thr Gly Leu Pro Ala Gly Glu Tyr Cys
                405                 410                 415
Asn Ile Leu Ala Gly Asp Ala Leu Cys Ser Gly Ser Thr Ile Lys Val
            420                 425                 430
Asp Ala Ser Gly Met Ala Thr Phe Asn Val Ala Gly Met Lys Ala Ala
        435                 440                 445
Ala Ile His Ile Asn Ala Lys Pro Asp Ser Thr Ser Ser Gly Ser Ser
    450                 455                 460
Gly Ser Ser Ser Gly Ser Ser Ser Ser Ala Thr Ser Asn Lys Phe Ala
465                 470                 475                 480
Ser Met Asn Leu Arg Gly Thr Asn Asn Gly Trp Ala Ser Thr Ala Met
                485                 490                 495
Thr Val Asp Ala Asn Arg Val Trp Ser Ala Asp Val Thr Phe Thr Gly
            500                 505                 510
Ala Ala Asp Ala Asn Gly Ala Gln Arg Phe Lys Phe Asp Val Tyr Gly
        515                 520                 525
Asn Trp Thr Glu Ser Tyr Gly Asp Thr Gln Ala Asp Gly Ile Ala Asp
    530                 535                 540
Lys Gly Ser Ala Lys Asp Ile Tyr Phe Asn Gly Val Gly Lys Tyr Arg
545                 550                 555                 560
Val Ser Leu Lys Glu Ser Asp Met Ser Tyr Thr Leu Thr Gln Leu Ser
                565                 570                 575
Ser Asn Gln Ala Pro Val Ala Ala Ile Thr Pro Lys Thr Leu Ser Val
            580                 585                 590
Lys Leu Gly Asp Ser Val Val Phe Asp Ala Ser Gly Ser Thr Asp Asp
```

595 600 605
Val Gly Val Thr Gly Tyr Ser Trp Ser Thr Gly Gly Ser Ala Lys Thr
610 615 620
Glu Thr Val Leu Phe Asp Ala Leu Gly Thr Lys Thr Ile Thr Val Thr
625 630 635 640
Val Ala Asp Ala Asp Gly Leu Thr Ser Lys Ala Ser Ala Thr Val Thr
645 650 655
Val Thr Asp Gly Ser Val Ala Tyr Asn Ser Asn Phe Ala Ser Leu Asn
660 665 670
Phe Arg Gly Thr Pro Asn Ser Trp Gly Ala Ala Ala Met Thr Leu Val
675 680 685
Ala Asp Asn Thr Trp Glu Ala Thr Val Asn Phe Asp Gly Gln Ala Asn
690 695 700
Gln Arg Phe Lys Phe Asp Ile Lys Gly Asp Trp Ser Gln Asn Tyr Gly
705 710 715 720
Asp Ser Asn Lys Asp Gly Val Ala Glu Arg Thr Gly Ala Asp Ile Tyr
725 730 735
Thr Thr Val Thr Gly Gln Tyr Lys Val Gln Phe Asn Asp Ser Thr Leu
740 745 750
Lys Tyr Thr Leu Thr Lys Leu Ala Asp Ser Ser Ala Thr Ser Tyr Ser
755 760 765
Ala Asn Phe Ala Ser Leu Tyr Leu Arg Gly Thr Pro Asn Ser Trp Gly
770 775 780
Thr Thr Ala Met Lys Leu Val Ala Asn Asn Ser Trp Gln Ala Glu Val
785 790 795 800
Thr Phe Thr Gly Lys Gly Asp Ala Thr Gly Ala Gln Arg Phe Lys Phe
805 810 815
Asp Val Lys Gly Asp Trp Ser Gln Asn Tyr Gly Asp Ser Asn Met Asp
820 825 830
Gly Thr Ala Glu Arg Thr Gly Gly Asp Ile Thr Ser Ala Val Val Gly
835 840 845
Thr Tyr Leu Val Thr Phe Asn Asp Ser Thr Leu Lys Tyr Thr Leu Thr
850 855 860
Ala Lys
865

<210> 205
<211> 1674
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 205
atgtaccgcg taatacctat tattttgatt atgagtatga ttgtagcttg tgagtctcca 60
aagaaaaaaa caaccgaaac cgctcaacct tcaacaaatg ccgaaaaacc ctttgtttgg 120
gaggctgcca atgtatattt tttgttaact gaccgtttta acaacggtaa cccaaacaat 180
gacatcaatt ttaataggac taaagaatca ggaaaactcc gcaattttat gggaggcgat 240
atcaagggca tcacccaaaa aataaatgag gggtatttta gtaaactagg cgttaatgcc 300
atctggctta ccccggttgt tgaacaaata catggcagtg ttgatgaagg taccggcaat 360
acctatgcct ttcatggcta ttgggccaaa gattggacaa acttagaccc aaattttggc 420
acaaaagaag accttgccga actggtggca actgcccatg caaaaggcat caggatactt 480
ttagatgtgg taataaacca caccggcccg gtaaccgacc aagacccggt ttggggagaa 540
gattgggtac gtacaggccc gcagtgtacc tatgataatt acaccaatac caccagttgc 600
acgctggtag ccaatttacc tgatatactt acagaaagta atgaaaatgt ggccttacca 660
acctttttgt tagataaatg gaaagccgaa ggcagattag agcaagaact aaaagaactt 720
gacgattttt tttcccgcac aggccaccca cgcgcacccc gcttttacat tattaaatgg 780
cttaccgatt acatccgaga atttggggta gatgggttta gggttgatac cgtaaaacat 840
accgaagaaa cggtttgggc cgagttgtat gatgaagccg taattgcttt tgccgaatat 900
aaaaaagcca acccagacaa ggtattggac gataatgaat tttatatggt aggcgaagtg 960
tacaactacg gtatttccgg cggaaggttc tatgatttcg gcgataaaaa ggtggactat 1020
tttgaccacg gatttaaaag cctcatcaat tttgaaatga aatatgatgc caattttacc 1080
tacgatacac tttttaggaa gtacgatacc cttttgcata ccaaacttaa aggcagaagt 1140
gtgctcaact acctctcatc tcacgacgat ggaagtccat ttgataaaat gcggcaaaaa 1200
ccatacgagt cggctacaaa attactgctc actccgggcg catcccaaat ttattacggt 1260
gacgaaaccg ccagaagcct taacatagaa ggcgcacagg gagatgctac gcttcgttcg 1320
tttatgaatt gggaagagct cgcagaagac cctgccaagc aaaaaatact tcagcattgg 1380

caaaaactgg gcagtttcag gaacaaccac cccgcagttg gtgccggaag gcacaaaacc 1440
cttggcaaaa agccgtttta cacctttagc agggtttatc aaaaaaatgg ttttattgac 1500
aaagttgtgg tagcattaga tgcccctaaa ggccaaaaac aaattaccgt taatggtgtt 1560
tttgatgacg gtacaaaact tgtagatgcc tattcaggca aagaaacctc agttaaaaat 1620
ggtatcgttt cactttcttc tgaatttgat attgttttgt tagaacaaaa ataa 1674

<210> 206
<211> 557
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 206
Met Tyr Arg Val Ile Pro Ile Ile Leu Ile Met Ser Met Ile Val Ala
 1   5   10   15
Cys Glu Ser Pro Lys Lys Lys Thr Thr Glu Thr Ala Gln Pro Ser Thr
            20   25   30
Asn Ala Glu Lys Pro Phe Val Trp Glu Ala Ala Asn Val Tyr Phe Leu
        35   40   45
Leu Thr Asp Arg Phe Asn Asn Gly Asn Pro Asn Asn Asp Ile Asn Phe
    50   55   60
Asn Arg Thr Lys Glu Ser Gly Lys Leu Arg Asn Phe Met Gly Gly Asp
65   70   75   80
Ile Lys Gly Ile Thr Gln Lys Ile Asn Glu Gly Tyr Phe Ser Lys Leu
                85   90   95
Gly Val Asn Ala Ile Trp Leu Thr Pro Val Val Glu Gln Ile His Gly
            100   105   110
Ser Val Asp Glu Gly Thr Gly Asn Thr Tyr Ala Phe His Gly Tyr Trp
        115   120   125
Ala Lys Asp Trp Thr Asn Leu Asp Pro Asn Phe Gly Thr Lys Glu Asp
    130   135   140
Leu Ala Glu Leu Val Ala Thr Ala His Ala Lys Gly Ile Arg Ile Leu
145   150   155   160
Leu Asp Val Val Ile Asn His Thr Gly Pro Val Thr Asp Gln Asp Pro
                165   170   175
Val Trp Gly Glu Asp Trp Val Arg Thr Gly Pro Gln Cys Thr Tyr Asp
            180   185   190
Asn Tyr Thr Asn Thr Thr Ser Cys Thr Leu Val Ala Asn Leu Pro Asp
        195   200   205
Ile Leu Thr Glu Ser Asn Glu Asn Val Ala Leu Pro Thr Phe Leu Leu
    210   215   220
Asp Lys Trp Lys Ala Glu Gly Arg Leu Glu Gln Glu Leu Lys Glu Leu
225   230   235   240
Asp Asp Phe Phe Ser Arg Thr Gly His Pro Arg Ala Pro Arg Phe Tyr
                245   250   255
Ile Ile Lys Trp Leu Thr Asp Tyr Ile Arg Glu Phe Gly Val Asp Gly
            260   265   270
Phe Arg Val Asp Thr Val Lys His Thr Glu Glu Thr Val Trp Ala Glu
        275   280   285
Leu Tyr Asp Glu Ala Val Ile Ala Phe Ala Glu Tyr Lys Lys Ala Asn
    290   295   300
Pro Asp Lys Val Leu Asp Asp Asn Glu Phe Tyr Met Val Gly Glu Val
305   310   315   320
Tyr Asn Tyr Gly Ile Ser Gly Gly Arg Phe Tyr Asp Phe Gly Asp Lys
                325   330   335
Lys Val Asp Tyr Phe Asp His Gly Phe Lys Ser Leu Ile Asn Phe Glu
            340   345   350
Met Lys Tyr Asp Ala Asn Phe Thr Tyr Asp Thr Leu Phe Arg Lys Tyr
        355   360   365
Asp Thr Leu Leu His Thr Lys Leu Lys Gly Arg Ser Val Leu Asn Tyr
    370   375   380
Leu Ser Ser His Asp Asp Gly Ser Pro Phe Asp Lys Met Arg Gln Lys
385   390   395   400
Pro Tyr Glu Ser Ala Thr Lys Leu Leu Leu Thr Pro Gly Ala Ser Gln
                405   410   415
Ile Tyr Tyr Gly Asp Glu Thr Ala Arg Ser Leu Asn Ile Glu Gly Ala

```
            420                 425                 430
Gln Gly Asp Ala Thr Leu Arg Ser Phe Met Asn Trp Glu Glu Leu Ala
        435                 440                 445
Glu Asp Pro Ala Lys Gln Lys Ile Leu Gln His Trp Gln Lys Leu Gly
    450                 455                 460
Ser Phe Arg Asn Asn His Pro Ala Val Gly Ala Gly Arg His Lys Thr
465                 470                 475                 480
Leu Gly Lys Lys Pro Phe Tyr Thr Phe Ser Arg Val Tyr Gln Lys Asn
                485                 490                 495
Gly Phe Ile Asp Lys Val Val Val Ala Leu Asp Ala Pro Lys Gly Gln
            500                 505                 510
Lys Gln Ile Thr Val Asn Gly Val Phe Asp Asp Gly Thr Lys Leu Val
        515                 520                 525
Asp Ala Tyr Ser Gly Lys Glu Thr Ser Val Lys Asn Gly Ile Val Ser
    530                 535                 540
Leu Ser Ser Glu Phe Asp Ile Val Leu Leu Glu Gln Lys
545                 550                 555
```

<210> 207
<211> 1378
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 207

```
ctgtcgactg agcctttcgt tttgggctcg agactgactc tcagcccacc ccgcagtagc      60
tccagacgga gtagccgtaa tagccgttgg ccgggtcgtg ggcaggggcc tcgaggtaca     120
cccacccgct tgagtccacc cacttgtcca cccagccgcc gaggttgccg gtgtactcgt     180
ggatgcacgc tcccgcgaac ttcggaacgt agacccacct tccggctttg cttgaggcga     240
ggttgatgta tgttatcagt cccggcttgc ttccgtagcc gtttctcacg aatatcagct     300
cgtcgttgtc gtagtaaacg acgtcagtgc ttcctccggc caggttgtca tgtatccaga     360
tgaggttctt gagcttatcc ttgttgagcc actcctcgta gtcgcggtag aatattgtcg     420
gctggccctc gtaggtgagg atgaacgcgt aggctggata cttgttccag attatatcgg     480
tgtcgtggtt tgcaacgaag gttacggcct taaacgggtc gcggctgacg actgtgcccc     540
cgttcttgag ggcctcgacg agtgcgggaa tgttcttgtt gtcaaaggcc gcgtccatct     600
tgtagtagag cgggaagtcg aagaccttgg cgccgctcga gtaggcccag ttgaggagtg     660
catcaacgtt ggtgtcccag tactcgccaa cggcccagcc gccccaccag ttgagccagt     720
ccttgacgac ccacgctccg tggcccttca cgtagtcaaa gcgccaggca tcaacgccga     780
tgctccttag gtaggcggcg tagctctcat cgctcgccca gagccagtgc tggtcccagc     840
tcttctcgtg ggctatgtct gggaagcctc caaatgtgcc ctcgtcacag cacttgacct     900
cgttggggtg gaagtcgagg tagttggcag tatatttgcc cgaggccacc tttgagaagt     960
ccgtccaggt gtagtcccca acgaacgggt tccactcgag gtctccgcct gcgcggtggt    1020
ttatgacgat gtccgctatg acctttatgc cgtaggcatg ggccgtgttt atcatgttca    1080
cgagctcctg cttggagcca aagcgcgtct ctaccgttcc cttctggtcg tactcaccga    1140
ggtcaaagaa gtcgtagggg tcgtagccca tcgaataggc gccgcccatg cccttgctcg    1200
ccgggggaat ccaaatggcg gatattcccg cctcgtacca ctccggtatc ttgctcctga    1260
tggtgtccca ccagattcct ccacctggga cgtcccagta gaaggcctgc attataacgc    1320
cgccctcttc cagctcggag tacttggcca taagttacct cctactagta gattaaaa      1378
```

<210> 208
<211> 439
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 208

```
Leu Ser Thr Glu Pro Phe Val Leu Gly Ser Arg Leu Thr Leu Ser Pro
 1               5                  10                  15
Pro Arg Ser Ser Ser Arg Arg Ser Ser Arg Asn Ser Arg Trp Pro Gly
            20                  25                  30
Arg Gly Gln Gly Pro Arg Gly Thr Pro Thr Arg Leu Ser Pro Pro Thr
        35                  40                  45
Cys Pro Pro Ser Arg Arg Gly Cys Arg Cys Thr Arg Gly Cys Thr Leu
    50                  55                  60
```

```
Pro Arg Thr Ser Glu Arg Arg Pro Thr Phe Arg Leu Cys Leu Arg Arg
65                  70                  75                  80
Gly Cys Met Leu Ser Val Pro Ala Cys Phe Arg Ser Arg Phe Ser Arg
                85                  90                  95
Ile Ser Ala Arg Arg Cys Arg Ser Lys Arg Arg Gln Cys Phe Leu Arg
            100                 105                 110
Pro Gly Cys His Val Ser Arg Gly Ser Ala Tyr Pro Cys Ala Thr Pro
        115                 120                 125
Arg Ser Arg Gly Arg Ile Leu Ser Ala Gly Pro Arg Arg Gly Thr Arg
    130                 135                 140
Arg Leu Asp Thr Cys Ser Arg Leu Tyr Arg Cys Arg Gly Leu Gln Arg
145                 150                 155                 160
Arg Leu Arg Pro Thr Gly Arg Gly Arg Leu Cys Pro Arg Ser Gly Pro
                165                 170                 175
Arg Arg Val Arg Glu Cys Ser Cys Cys Gln Arg Pro Arg Pro Ser Cys
            180                 185                 190
Ser Arg Ala Gly Ser Arg Arg Pro Trp Arg Arg Ser Ser Arg Pro Ser
        195                 200                 205
Gly Val His Gln Arg Trp Cys Pro Ser Thr Arg Gln Arg Pro Ser Arg
    210                 215                 220
Pro Thr Ser Ala Ser Pro Arg Pro Thr Leu Arg Gly Pro Ser Arg Ser
225                 230                 235                 240
Gln Ser Ala Arg His Gln Arg Arg Cys Ser Leu Gly Arg Arg Arg Ser
                245                 250                 255
Ser His Arg Ser Pro Arg Ala Ser Ala Gly Pro Ser Ser Ser Arg Gly
            260                 265                 270
Leu Cys Leu Gly Ser Leu Gln Met Cys Pro Arg His Ser Thr Pro Arg
        275                 280                 285
Trp Gly Gly Ser Arg Gly Ser Trp Gln Tyr Ile Cys Pro Arg Pro Pro
    290                 295                 300
Leu Arg Ser Pro Ser Arg Cys Ser Pro Gln Arg Thr Gly Ser Thr Arg
305                 310                 315                 320
Gly Leu Arg Leu Arg Gly Gly Leu Arg Cys Pro Leu Pro Leu Cys Arg
                325                 330                 335
Arg His Gly Pro Cys Leu Ser Cys Ser Arg Ala Pro Ala Trp Ser Gln
            340                 345                 350
Ser Ala Ser Leu Pro Phe Pro Ser Gly Arg Thr His Arg Gly Gln Arg
        355                 360                 365
Ser Arg Arg Gly Arg Ser Pro Ser Asn Arg Arg Arg Pro Cys Pro Cys
    370                 375                 380
Ser Pro Gly Glu Ser Lys Trp Arg Ile Phe Pro Pro Arg Thr Thr Pro
385                 390                 395                 400
Val Ser Cys Ser Trp Cys Pro Thr Arg Phe Leu His Leu Gly Arg Pro
                405                 410                 415
Ser Arg Arg Pro Ala Leu Arg Arg Pro Leu Pro Ala Arg Ser Thr Trp
            420                 425                 430
Pro Val Thr Ser Tyr Ile Lys
        435
```

<210> 209
<211> 1416
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 209

```
atgattcagc ccatgcactc tcgggaacag gcctgccgtc tcattccggc actgatcatg      60
acatttgcac tggcactgcc gttgcaaatt cgtgccgatg tcaccctgca tgctttcaac     120
tggagctatg ccgatgtcgc tgatcgggcc gttgacatcg ctgcagcagg gtacagtgcc     180
gtgctggtgg ccccgccact tcgatccgaa ggcacggcct ggtgggcgcg ataccagccc     240
caggatctcc gccttatcga ccatccgctg ggcaatacac atgacttcgt caacatgatc     300
gatgctctcg atgatgtggg tgtgggcgtg tacgccgaca tcgtgctcaa ccacatggcc     360
aatgaggctg cacaaaggcc tgacctgaac taccctggtc aggcagtgct tgacgaatat     420
gcttccgatc ccggtcattt cgagggcttg aggctgttcg gtaatctgag cttcaatttc     480
ctgtcggaac atgatttcgg acccgcccag tgcattcagg attacagcga tgtgtttcag     540
gtccagaact ggcggctgtg cggaccgccg ccggacccgg gcctgcccga cctggtcgcc     600
```

aatgactggg tgatctctca acagcgccag tatctggaag ccatcaaggc gctgggtgtg 660
gctggcatgc gcatcgacgc ggtcaagcat atgcccatga gccatatcaa tgccgttctc 720
acccccgaga tccggtcggg cttgcatgtg tttggcgaag tcatcacctc cggtggggct 780
ggtgatacat cctacgaccg ttttctggcc ccttacctgg cacaaagcga ccatggtgcc 840
tatgactttc cattgtttga aaccattcgc cgtgctttcg gcttcggtgg cagcatgagt 900
gaactggtcg atcctgctgc ctacggtcag gccctgccac cggaccgcgc catcaccttc 960
gtcatcacgc acgatattcc gaacaatgac ggatttcgct accagatact cgaccccgtc 1020
gatgaatcac tggcctacgc ctacattctg ggccgcgatg gcggtgtccc gcttctgtat 1080
tccgacaaca atgaaagcgg cgatggccgc tggatcgatg cctggcaacg tccggatctg 1140
gttgcaatgg tcggcttcca caatgcagtc cacggtcagg acatggccgt gctttcacat 1200
gacgactgcc acctgctgtt tcggcgcggc agcctcggga ttgtcggcat caacaagtgc 1260
ggccatgcac tcagctcctg ggtcaacatg aaccagagcg tactgtggtg gtacgcggac 1320
tacacagacg tgctcgacag caacagcgtt gtcaacatcc agtcatcctg gcacgagttc 1380
atccttcccg cccgccaggc acgcctgtgg ttgcga 1416

<210> 210
<211> 472
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 210
Met Ile Gln Pro Met His Ser Arg Glu Gln Ala Cys Arg Leu Ile Pro
 1   5   10   15
Ala Leu Ile Met Thr Phe Ala Leu Ala Leu Pro Leu Gln Ile Arg Ala
            20   25   30
Asp Val Thr Leu His Ala Phe Asn Trp Ser Tyr Ala Asp Val Ala Asp
        35   40   45
Arg Ala Val Asp Ile Ala Ala Ala Gly Tyr Ser Ala Val Leu Val Ala
    50   55   60
Pro Pro Leu Arg Ser Glu Gly Thr Ala Trp Trp Ala Arg Tyr Gln Pro
65   70   75   80
Gln Asp Leu Arg Leu Ile Asp His Pro Leu Gly Asn Thr His Asp Phe
                85   90   95
Val Asn Met Ile Asp Ala Leu Asp Asp Val Gly Val Gly Val Tyr Ala
            100   105   110
Asp Ile Val Leu Asn His Met Ala Asn Glu Ala Ala Gln Arg Pro Asp
        115   120   125
Leu Asn Tyr Pro Gly Gln Ala Val Leu Asp Glu Tyr Ala Ser Asp Pro
    130   135   140
Gly His Phe Glu Gly Leu Arg Leu Phe Gly Asn Leu Ser Phe Asn Phe
145   150   155   160
Leu Ser Glu His Asp Phe Gly Pro Ala Gln Cys Ile Gln Asp Tyr Ser
                165   170   175
Asp Val Phe Gln Val Gln Asn Trp Arg Leu Cys Gly Pro Pro Pro Asp
            180   185   190
Pro Gly Leu Pro Asp Leu Val Ala Asn Asp Trp Val Ile Ser Gln Gln
        195   200   205
Arg Gln Tyr Leu Glu Ala Ile Lys Ala Leu Gly Val Ala Gly Met Arg
    210   215   220
Ile Asp Ala Val Lys His Met Pro Met Ser His Ile Asn Ala Val Leu
225   230   235   240
Thr Pro Glu Ile Arg Ser Gly Leu His Val Phe Gly Glu Val Ile Thr
                245   250   255
Ser Gly Gly Ala Gly Asp Thr Ser Tyr Asp Arg Phe Leu Ala Pro Tyr
            260   265   270
Leu Ala Gln Ser Asp His Gly Ala Tyr Asp Phe Pro Leu Phe Glu Thr
        275   280   285
Ile Arg Arg Ala Phe Gly Phe Gly Gly Ser Met Ser Glu Leu Val Asp
    290   295   300
Pro Ala Ala Tyr Gly Gln Ala Leu Pro Pro Asp Arg Ala Ile Thr Phe
305   310   315   320
Val Ile Thr His Asp Ile Pro Asn Asn Asp Gly Phe Arg Tyr Gln Ile
                325   330   335
Leu Asp Pro Val Asp Glu Ser Leu Ala Tyr Ala Tyr Ile Leu Gly Arg
            340   345   350

```
Asp Gly Gly Val Pro Leu Leu Tyr Ser Asp Asn Asn Glu Ser Gly Asp
        355                 360                 365
Gly Arg Trp Ile Asp Ala Trp Gln Arg Pro Asp Leu Val Ala Met Val
    370                 375                 380
Gly Phe His Asn Ala Val His Gly Gln Asp Met Ala Val Leu Ser His
385                 390                 395                 400
Asp Asp Cys His Leu Leu Phe Arg Arg Gly Ser Leu Gly Ile Val Gly
                405                 410                 415
Ile Asn Lys Cys Gly His Ala Leu Ser Ser Trp Val Asn Met Asn Gln
            420                 425                 430
Ser Val Leu Trp Trp Tyr Ala Asp Tyr Thr Asp Val Leu Asp Ser Asn
        435                 440                 445
Ser Val Val Asn Ile Gln Ser Ser Trp His Glu Phe Ile Leu Pro Ala
    450                 455                 460
Arg Gln Ala Arg Leu Trp Leu Arg
465                 470
```

<210> 211
<211> 1491
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 211

```
gtgtttcgtt ctgacacagt ttcgcgtacc tgcatgtatg gtgcgctgcg taatgcctac      60
caacccgatc gggtgtttac tggagtcacg gtgcggacat gcaacttaaa aaagcatgct     120
catcgccagg cgctgttgtt catcgtgacg cggtgcctgt gcctgaaatc caggcagacc     180
cataaaaaca acaacaaacc gataacaaac gacccaagcc ttctaagagg agaaaacggg     240
atggctttta aactacgcaa aaaggcgctc gttggcctgt tcacggccgg cgcaatggta     300
tatgccggtg cagcggcgag tggtgaaatc attctgcagg gcttccactg gcactccaag     360
tggggcggca acaatcaggg ttggtggcag gtgatggaag gtcaggccaa caccatcgcc     420
aacgccggct ttacgcacgt gtggttcccg ccggtccata actcggccga tgccgagggt     480
tacctacccc gcgagctgaa caacctcaac tccagctatg gctccgaagc acagctgcgc     540
agcgccatcc aggcactgaa caatcgcggc gtgcatgcga ttgccgatgt ggtcatgaac     600
caccgggtgg gctgctctgg ctgggcggat ttctgtaacc cggactggcc gacctggtac     660
atcgtcgcca atgattcctg gcccggtggc ccgaaaagcc agaactggga cacgggtgag     720
acgtaccacg ccgcccgtga cctcgatcac gccaatccgc aggtgcgcaa cgatatctcg     780
cactacctga acagccgcct caaggacgtc ggcttctccg gctggcgctg ggactatgcc     840
aagggtttct ggcccggcta tgtcggcgag tacaactgga acaccaaccc gaacttctgt     900
gtgggtgagg tgtgggacga tctcgacccc aacaatccca acccgcaccg ccagcaactg     960
gtggactggg ttgatgctac cggtggcagt tgtcacgtct tcgacttcac caccaagggg    1020
ctgacgaact atgcgctgca gcatggccag tactggcgcc tgcagggtga taatggtggc    1080
ccggctggcg gcatcggctg gtggccgcaa cgcatggtga ccttcgtcga caaccatgac    1140
acgggcccga gcaatcactg tggtgacggc cagaacctct ggcccgtgcc ctgtgacaag    1200
gtcatggagg cgtatgccta catcctgacc catccgggcg tgccgtcggt gtactggacg    1260
cacttcttca actggaatct tggtagcgag atcagccagt tgatgcagat ccgcaagaac    1320
cagggcgtgc actccggttc cgacgtctgg atcgccgagg cccgtcacgg cctgtacgcc    1380
gcctatatca acggtaatgt ggcgatgaag atgggctggg ataactggag cccgggctgg    1440
ggctggtcgc tggcggcctc cggtaacaac tgggccgtct ggacacgctg a              1491
```

<210> 212
<211> 496
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 212

```
Val Phe Arg Ser Asp Thr Val Ser Arg Thr Cys Met Tyr Gly Ala Leu
 1               5                  10                  15
Arg Asn Ala Tyr Gln Pro Asp Arg Val Phe Thr Gly Val Thr Val Arg
            20                  25                  30
Thr Cys Asn Leu Lys Lys His Ala His Arg Gln Ala Leu Leu Phe Ile
        35                  40                  45
Val Thr Arg Cys Leu Cys Leu Lys Ser Arg Gln Thr His Lys Asn Asn
```

```
    50                  55                  60
Asn Lys Pro Ile Thr Asn Asp Pro Ser Leu Leu Arg Gly Glu Asn Gly
65                  70                  75                  80
Met Ala Phe Lys Leu Arg Lys Lys Ala Leu Val Gly Leu Phe Thr Ala
                85                  90                  95
Gly Ala Met Val Tyr Ala Gly Ala Ala Ala Ser Gly Glu Ile Ile Leu
            100                 105                 110
Gln Gly Phe His Trp His Ser Lys Trp Gly Gly Asn Asn Gln Gly Trp
        115                 120                 125
Trp Gln Val Met Glu Gly Gln Ala Asn Thr Ile Ala Asn Ala Gly Phe
    130                 135                 140
Thr His Val Trp Phe Pro Pro Val His Asn Ser Ala Asp Ala Glu Gly
145                 150                 155                 160
Tyr Leu Pro Arg Glu Leu Asn Asn Leu Asn Ser Ser Tyr Gly Ser Glu
                165                 170                 175
Ala Gln Leu Arg Ser Ala Ile Gln Ala Leu Asn Asn Arg Gly Val His
            180                 185                 190
Ala Ile Ala Asp Val Val Met Asn His Arg Val Gly Cys Ser Gly Trp
        195                 200                 205
Ala Asp Phe Cys Asn Pro Asp Trp Pro Thr Trp Tyr Ile Val Ala Asn
    210                 215                 220
Asp Ser Trp Pro Gly Gly Pro Lys Ser Gln Asn Trp Asp Thr Gly Glu
225                 230                 235                 240
Thr Tyr His Ala Ala Arg Asp Leu Asp His Ala Asn Pro Gln Val Arg
                245                 250                 255
Asn Asp Ile Ser His Tyr Leu Asn Ser Arg Leu Lys Asp Val Gly Phe
            260                 265                 270
Ser Gly Trp Arg Trp Asp Tyr Ala Lys Gly Phe Trp Pro Gly Tyr Val
        275                 280                 285
Gly Glu Tyr Asn Trp Asn Thr Asn Pro Asn Phe Cys Val Gly Glu Val
    290                 295                 300
Trp Asp Asp Leu Asp Pro Asn Asn Pro Asn Pro His Arg Gln Gln Leu
305                 310                 315                 320
Val Asp Trp Val Asp Ala Thr Gly Gly Ser Cys His Val Phe Asp Phe
                325                 330                 335
Thr Thr Lys Gly Leu Thr Asn Tyr Ala Leu Gln His Gly Gln Tyr Trp
            340                 345                 350
Arg Leu Gln Gly Asp Asn Gly Gly Pro Ala Gly Gly Ile Gly Trp Trp
        355                 360                 365
Pro Gln Arg Met Val Thr Phe Val Asp Asn His Asp Thr Gly Pro Ser
    370                 375                 380
Asn His Cys Gly Asp Gly Gln Asn Leu Trp Pro Val Pro Cys Asp Lys
385                 390                 395                 400
Val Met Glu Ala Tyr Ala Tyr Ile Leu Thr His Pro Gly Val Pro Ser
                405                 410                 415
Val Tyr Trp Thr His Phe Phe Asn Trp Asn Leu Gly Ser Glu Ile Ser
            420                 425                 430
Gln Leu Met Gln Ile Arg Lys Asn Gln Gly Val His Ser Gly Ser Asp
        435                 440                 445
Val Trp Ile Ala Glu Ala Arg His Gly Leu Tyr Ala Ala Tyr Ile Asn
    450                 455                 460
Gly Asn Val Ala Met Lys Met Gly Trp Asp Asn Trp Ser Pro Gly Trp
465                 470                 475                 480
Gly Trp Ser Leu Ala Ala Ser Gly Asn Asn Trp Ala Val Trp Thr Arg
                485                 490                 495

<210> 213
<211> 23
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 213
Met Phe Leu Leu Ala Phe Leu Leu Thr Ala Ser Leu Phe Cys Pro Thr
 1               5                  10                  15
Gly Gln Pro Ala Lys Ala Ala
```

20

<210> 214
<211> 23
<212> PRT
<213> Bacterial

<400> 214
Val Leu Thr Phe His Arg Ile Ile Arg Lys Gly Trp Met Phe Leu Leu
 1   5   10   15
Ala Phe Leu Leu Thr Ala Ser
            20

<210> 215
<211> 33
<212> PRT
<213> Bacterial

<400> 215
Met Lys Ser Phe Ala Phe Met Pro Ile Leu Phe Tyr Ala Asn Asp Phe
 1   5   10   15
Ile Ser Glu Arg Glu Gly Gly Gly Lys Met Gly Lys Asn Met Arg Arg
            20   25   30
Arg

<210> 216
<211> 31
<212> PRT
<213> Bacterial

<400> 216
Met Arg Lys Lys Met Ser His Ser Arg Phe Thr Phe Leu Leu Ile Leu
 1   5   10   15
Ala Leu Phe Ile Phe Phe Ser Gly Cys Ile Ser Glu Val Lys Ser
            20   25   30

<210> 217
<211> 30
<212> PRT
<213> Bacterial

<400> 217
Met Tyr Thr Leu Phe Ile Arg Ser Phe Tyr Asp Thr Asn Asn Asp Gly
 1   5   10   15
Val Gly Asp Tyr Asn Gly Val Ala Gln Lys Val Asp Tyr Leu
            20   25   30

<210> 218
<211> 22
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 218
Val Leu Thr Phe His Arg Ile Ile Arg Lys Gly Trp Met Phe Leu Leu
 1   5   10   15
Ala Phe Leu Leu Thr Ala
            20

<210> 219
<211> 33
<212> PRT
<213> Unknown

<220>

<223> Obtained from an environmental sample.

<400> 219
Met Ser Leu Phe Lys Lys Ile Phe Pro Trp Ile Val Ser Leu Leu Leu
 1   5   10   15
Leu Phe Ser Phe Ile Ala Pro Phe Ser Ile Gln Thr Glu Lys Val Arg
            20   25   30
Ala

<210> 220
<211> 25
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 220
Met Ala Arg Lys Thr Leu Ala Ile Phe Phe Val Leu Leu Val Leu Leu
 1   5   10   15
Ser Leu Ser Ala Val Pro Ala Lys Ala
            20   25

<210> 221
<211> 35
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 221
Met Pro Ala Phe Lys Ser Lys Val Met His Met Lys Leu Lys Tyr Leu
 1   5   10   15
Ala Leu Val Leu Leu Ala Val Ala Ser Ile Gly Leu Leu Ser Thr Pro
            20   25   30
Val Gly Ala
        35

<210> 222
<211> 28
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 222
Met Lys Gln Gln Lys Arg Leu Tyr Ala Arg Leu Leu Thr Leu Leu Phe
 1   5   10   15
Ala Leu Ile Phe Leu Leu Pro His Ser Ala Ala Ala
            20   25

<210> 223
<211> 21
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 223
Met Arg Val Phe Leu Val Val Pro Lys Leu Ser Arg Pro Phe Gln Ala
 1   5   10   15
Glu Ser Gln Gln Gln
            20

<210> 224
<211> 30
<212> PRT
<213> Bacterial

<400> 224
Met Gln Thr Phe Ala Phe Leu Phe Tyr Ser Lys Lys Gly Trp Val Cys
 1   5   10   15
Met Asn Tyr Leu Lys Lys Val Trp Leu Tyr Tyr Ala Ile Val
            20   25   30

<210> 225
<211> 35
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 225
Met Pro Gln Ala Ile Arg Thr Phe Ser Arg Trp Thr Leu Phe Gly Leu
 1   5   10   15
Ile Gly Val Phe Leu Leu Gly Leu Val Phe Ser Val Pro Pro Arg Ala
            20   25   30
Ile Gln Ala
        35

<210> 226
<211> 28
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 226
Val Val His Met Lys Leu Lys Tyr Leu Ala Leu Val Leu Leu Ala Val
 1   5   10   15
Ala Ser Ile Gly Leu Leu Ser Thr Pro Val Gly Ala
            20   25

<210> 227
<211> 30
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 227
Val Cys Met Asn Tyr Leu Lys Lys Val Trp Leu Tyr Tyr Ala Ile Val
 1   5   10   15
Ala Thr Leu Ile Ile Tyr Phe Leu Thr Pro Phe Ser Thr Ala
            20   25   30

<210> 228
<211> 30
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 228
Met Pro Gln Leu Tyr Pro Leu Pro Pro Arg Trp Arg Arg Ala Ala Arg
 1   5   10   15
Gln Gly Leu Ala Ala Leu Thr Leu Ala Thr Thr Ala Leu Gly
            20   25   30

<210> 229
<211> 30
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 229
Met Asn Asn Val Lys Lys Val Trp Leu Tyr Tyr Ser Ile Ile Ala Thr
 1   5   10   15
Leu Val Ile Ser Phe Phe Thr Pro Phe Ser Thr Ala Gln Ala
            20   25   30

<210> 230
<211> 28
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 230
Val Gly Arg Ala Gly Leu Ala His His Ser Asn Thr Ser Ala Lys Gly
 1   5   10   15
Thr Tyr Gly Ser Pro Leu Glu Leu Arg Pro Asp Arg
            20   25

<210> 231
<211> 23
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 231
Met Lys Thr Phe Asn Leu Lys Pro Thr Leu Leu Pro Leu Thr Leu Leu
 1   5   10   15
Leu Ser Ser Pro Val Leu Ala
            20

<210> 232
<211> 23
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 232
Met Lys Pro Ile Asn Thr Leu Leu Ile Ser Ala Leu Ala Val Cys Ser
 1   5   10   15
Phe Ser Ser Ala Thr Tyr Ala
            20

<210> 233
<211> 30
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 233
Met Pro Lys Ser Thr Phe Thr Lys Ser Ile Thr Lys Ser Leu Leu Ala
 1   5   10   15

Thr Ser Val Val Val Ser Leu Leu Pro Ala Tyr Ala Gln Ala
            20   25   30

<210> 234
<211> 27
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 234
Met Leu Lys Arg Ile Thr Val Val Cys Leu Leu Phe Ile Leu Leu Phe
 1   5   10   15
Pro Asn Ile Tyr Gly Arg Asn Lys Ala Glu Ala
            20   25

<210> 235
<211> 29
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 235
Met Ser Leu Asn Asn Phe Lys Val Lys Leu Leu Ser Phe Ala Val Ser
 1   5   10   15
Ser Ala Val Leu Ser Leu Ala Pro Asn Leu Ala Asn Ala
            20   25

<210> 236
<211> 28
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 236
Met Ile Leu Ser Asn Phe Lys Val Lys Leu Leu Ser Phe Ala Val Ser
 1   5   10   15
Ser Ala Val Leu Thr Leu Ala Ala Asn Val Ala Asn
            20   25

<210> 237
<211> 27
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 237
Met Leu Lys Arg Ile Thr Val Val Cys Leu Leu Phe Ile Leu Leu Phe
 1   5   10   15
Pro Asn Ile Tyr Glu Gly Asn Lys Ala Glu Ala
            20   25

<210> 238
<211> 26
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 238

Met Pro Ser Ile Asn Ala Ser Asp Cys Lys Lys Lys Gly Asp Arg Ser
1 5 10 15
Met Lys Arg Lys Lys Trp Thr Ala Leu Ala
20 25

<210> 239
<211> 33
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 239
Val Ser Arg Met Phe Ala Lys Arg Phe Lys Thr Ser Leu Leu Pro Leu
1 5 10 15
Phe Ala Gly Phe Leu Leu Leu Phe His Leu Val Leu Ala Gly Pro Thr
20 25 30
Ala

<210> 240
<211> 25
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 240
Met Gln Thr Ile Ala Lys Lys Gly Asp Glu Thr Met Lys Gly Lys Lys
1 5 10 15
Trp Thr Ala Leu Ala Leu Thr Leu Pro
20 25

<210> 241
<211> 25
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 241
Met Gln Lys Lys Gly Asp Glu Thr Met Lys Gly Lys Lys Trp Thr Ala
1 5 10 15
Leu Ala Leu Thr Leu Pro Leu Ala Ala
20 25

<210> 242
<211> 36
<212> PRT
<213> Bacterial

<400> 242
Val Asp Pro Lys Asn Cys Ser Gln Phe Met Gln Thr Ile Ala Lys Lys
1 5 10 15
Gly Asp Glu Thr Met Lys Gly Lys Lys Trp Thr Ala Leu Ala Leu Thr
20 25 30
Leu Pro Leu Ala
35

<210> 243
<211> 36
<212> PRT
<213> Unknown

<220>

<223> Obtained from an environmental sample.

<400> 243
Met Gln Thr Ile Ala Lys Lys Gly Asp Glu Thr Met Lys Gly Lys Lys
 1   5   10   15
Trp Thr Ala Leu Ala Leu Thr Leu Pro Leu Ala Ala Ser Leu Ser Thr
            20   25   30
Gly Val His Ala
        35

<210> 244
<211> 23
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 244
Met Lys Thr Phe Lys Leu Lys Arg Thr Phe Leu Pro Leu Thr Leu Leu
 1   5   10   15
Leu Ser Ala Pro Ala Phe Ala
            20

<210> 245
<211> 25
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 245
Met Gln Thr Ile Ala Lys Lys Gly Asp Glu Thr Met Lys Gly Lys Lys
 1   5   10   15
Trp Thr Ala Leu Ala Leu Thr Leu Pro
            20   25

<210> 246
<211> 22
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 246
Met Lys Asn Ile Ile Arg Leu Cys Ala Ala Ser Ala Ile Leu Thr Val
 1   5   10   15
Ser His Ala Ser Tyr Ala
            20

<210> 247
<211> 23
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 247
Met Lys Thr Phe Lys Leu Lys Arg Thr Phe Leu Pro Leu Thr Leu Leu
 1   5   10   15
Leu Ser Ala Pro Ala Phe Ala
            20

<210> 248
<211> 23

<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 248
Met Lys Thr Phe Lys Leu Lys Arg Thr Phe Leu Pro Leu Thr Leu Leu
 1   5   10   15
Leu Ser Ala Pro Ala Phe Ala
            20

<210> 249
<211> 25
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 249
Met Lys Leu Met Lys Gly Lys Lys Trp Thr Ala Leu Ala Leu Thr Leu
 1   5   10   15
Pro Leu Ala Ala Ser Leu Ser Thr Gly
            20   25

<210> 250
<211> 36
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 250
Met Gln Thr Ile Ala Lys Lys Gly Asp Glu Thr Met Lys Gly Lys Lys
 1   5   10   15
Trp Thr Ala Leu Ala Leu Thr Leu Pro Leu Ala Ala Ser Leu Ser Thr
            20   25   30
Gly Val His Ala
        35

<210> 251
<211> 25
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 251
Met Gln Thr Ile Ala Lys Lys Gly Asp Glu Thr Met Lys Gly Lys Lys
 1   5   10   15
Trp Thr Ala Leu Ala Leu Thr Leu Pro
            20   25

<210> 252
<211> 25
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 252
Met Lys Phe Lys Lys Ser Leu Ser Ala Gly Leu Leu Leu Phe Gly Gly
 1   5   10   15
Leu Ser Gly Val Thr Pro Ser Val Ala

20 25

<210> 253
<211> 23
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 253
Met Lys Pro Ser Lys Phe Val Phe Leu Ser Ala Ala Ile Ala Cys Ser
1 5 10 15
Leu Ser Ser Thr Ala Asn Ala
20

<210> 254
<211> 23
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 254
Val Ser Leu Thr Lys Lys Ala Gln Tyr Glu Pro Asn Thr Ala Pro Arg
1 5 10 15
Leu Ser Thr Ser Leu Gln Ser
20

<210> 255
<211> 30
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 255
Met Thr Ala Lys Ala Asp Asp Leu Arg Ile Tyr Gln Ile Met Val Glu
1 5 10 15
Ser Phe Val Asp Gly Asp Lys Gln Val Gly His Gly Thr Gly
20 25 30

<210> 256
<211> 25
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 256
Met Lys Met Lys Ser Arg Ala Trp Leu Leu Gly Ser Ala Val Ala Met
1 5 10 15
Ala Leu Ala Ser Ser Ala Ala Asn Ala
20 25

<210> 257
<211> 16
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 257
Met Tyr Arg Val Ile Pro Ile Ile Leu Ile Met Ser Met Ile Val Ala

1 5 10 15

<210> 258
<211> 47
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 258
gaacactagt aggaggtaac ttatggcaaa gtattccgag ctcgaag 47

<210> 259
<211> 39
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 259
gaacggtctc attccgccag ccagcaaggg gatgagcgg 39

<210> 260
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 260
gaaccgtctc aaaacacggc ccatgcctac ggc 33

<210> 261
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 261
gaacgtctca cctcgacttc caccccaacg aggtcaag 38

<210> 262
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 262
gaacgtctca ggcgctttga ctacgtgaag ggc 33

<210> 263
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 263
gaacggtctc aacaagatgg atgaggcctt tg 32

<210> 264
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 264
gaaccgtctc acgatataat ctggaacaag taccttgc 38

<210> 265
<211> 35
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 265
gaaccgtctc agaagcacga gcatagttta ctacg 35

<210> 266
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 266
gaaccgtctc aaaggtgggt ttatgtgccg 30

<210> 267
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 267
gaacgtctca ggaatccaaa tggcggatat tcccgc 36

<210> 268
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 268
gaacggtctc agtttatcat attgatgagc tcc 33

<210> 269
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 269
gaaccgtctc agaggtagtt ggcagtatat ttg 33

<210> 270
<211> 31

<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 270
gaacgtctca cgccaggcat caacgccgat g 31

<210> 271
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 271
gaacgtctca ttgtagtaga gcgggaagtc 30

<210> 272
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 272
gaacggtctc aatcggtgtc gtggtttgct ac 32

<210> 273
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 273
gaaccgtctc acttccacct gcgaggtggt c 31

<210> 274
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 274
gaaccgtctc accttccaac cttgctcgag c 31

<210> 275
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 275
tcgagactga ctctcaccca acaccgcaat agc 33

<210> 276
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 276
gaacactagt aggaggtaac ttatggccaa gtacctggag ctcgaagagg 50

<210> 277
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 277
gaacggtctc attcccccgg cgagcaaggg c 31

<210> 278
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 278
gaaccgtctc aaaacaccgc ccacgcctac gg 32

<210> 279
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 279
gaacgtctca cctcgacttc caccccaac 29

<210> 280
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 280
gaacgtctca ggcgcttcga ctacgtcaag g 31

<210> 281
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 281
gaacggtctc aacaagatgg acgcggcctt tgac 34

<210> 282
<211> 35
<212> DNA
<213> Artificial Sequence

<220>

<223> Primer

<400> 282
gaaccgtctc acgatataat ttggaacaag taccc 35

<210> 283
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 283
gaaccgtctc agaagcaccg acatagtcta c 31

<210> 284
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 284
gaaccgtctc aaaggtgggt ctacgttccg 30

<210> 285
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 285
gaacgtctca ggaatccata ttgcggagat tccggc 36

<210> 286
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 286
gaacggtctc agtttatcat gttcacgagc tc 32

<210> 287
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 287
gaaccgtctc agaggtagtt ggccgtgtac ttg 33

<210> 288
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 288
gaacgtctca gccatgcgtc aacgccgatg 30

<210> 289
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 289
gaacgtctca ttgtagtaga gcgggaagtc g 31

<210> 290
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 290
gaacggtctc aatcggtgtc gtggtttgca acg 33

<210> 291
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 291
gaaccgtctc acttccaccg gcgaggtggt cgtg 34

<210> 292
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 292
gaaccgtctc accttccggc cttgctcgag cc 32

<210> 293
<211> 35
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 293
tcgagactga ctctcagccc accccgcagt agctc 35

<210> 294
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 294
gaacactagt aggaggtaac ttatggccaa gtactccgag ctggaagagg 50

<210> 295
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 295
gaacggtctc attcctcccg cgagcaaggg 30

<210> 296
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 296
gaaccgtctc aaaacaccgc ccacgcctat g 31

<210> 297
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 297
gaacgtctca cctcgacttc cacccgaacg agc 33

<210> 298
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 298
gaacgtctca ggcgcttcga ctacgtcaag g 31

<210> 299
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 299
gaacggtctc aacaagatgg acgaggcctt cg 32

<210> 300
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 300
gaaccgtctc acgatataat ctggaacaag 30

<210> 301

<211> 35
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 301
gaaccgtctc agaagcactg acatcgttta ctacg 35

<210> 302
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 302
gaaccgtctc aaaggtgggt ttacgttccg 30

<210> 303
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 303
gaacgtctca ggaatccata tcgccgaaat 30

<210> 304
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 304
gaacggtctc agtttatcat gtttatgagc 30

<210> 305
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 305
gaaccgtctc agaggtagtt ggccgtgtat ttac 34

<210> 306
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 306
gaacgtctca cgccaggcat cgatgccgat 30

<210> 307
<211> 34
<212> DNA

<213> Artificial Sequence

<220>
<223> Primer

<400> 307
gaacgtctca ttgtagtaga gggcgaagtc aaag 34

<210> 308
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 308
gaacggtctc aatcggtatc gtggttggct acaaac 36

<210> 309
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 309
gaaccgtctc acttcctccg gcgaggttgt catg 34

<210> 310
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 310
gaaccgtctc accttccggc tttgcttgag gc 32

<210> 311
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 311
tcgagactga ctctcaccca acaccgcagt agctcc 36

<210> 312
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 312
cacacagcag caaccaacct cgagactgac tctcascc 38

<210> 313
<211> 461
<212> PRT
<213> Pyrococcus sp.

<220>

<400> 313

```
Met Lys Lys Phe Val Ala Leu Leu Ile Thr Met Phe Phe Val Val Ser
 1               5                  10                  15
Met Ala Ala Val Ala Gln Pro Ala Ser Ala Ala Lys Tyr Ser Glu Leu
            20                  25                  30
Glu Glu Gly Gly Val Ile Met Gln Ala Phe Tyr Trp Asp Val Pro Ala
        35                  40                  45
Gly Gly Ile Trp Trp Asp Thr Ile Arg Ser Lys Ile Pro Glu Trp Tyr
    50                  55                  60
Glu Ala Gly Ile Ser Ala Ile Trp Ile Pro Pro Ala Ser Lys Gly Met
65                  70                  75                  80
Gly Gly Ala Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Phe Phe Asp Leu
                85                  90                  95
Gly Glu Tyr Asn Gln Lys Gly Thr Val Glu Thr Arg Phe Gly Ser Lys
            100                 105                 110
Gln Glu Leu Ile Asn Met Ile Asn Thr Ala His Ala Tyr Gly Ile Lys
        115                 120                 125
Val Ile Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly Asp Leu Glu
    130                 135                 140
Trp Asn Pro Phe Val Gly Asp Tyr Thr Trp Thr Asp Phe Ser Lys Val
145                 150                 155                 160
Ala Ser Gly Lys Tyr Thr Ala Asn Tyr Leu Asp Phe His Pro Asn Glu
                165                 170                 175
Val Lys Cys Cys Asp Glu Gly Thr Phe Gly Gly Phe Pro Asp Ile Ala
            180                 185                 190
His Glu Lys Glu Trp Asp Gln His Trp Leu Trp Ala Ser Asp Glu Ser
        195                 200                 205
Tyr Ala Ala Tyr Leu Arg Ser Ile Gly Val Asp Ala Trp Arg Phe Asp
    210                 215                 220
Tyr Val Lys Gly Tyr Gly Ala Trp Val Val Lys Asp Trp Leu Asn Trp
225                 230                 235                 240
Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn Val Asp Ala
                245                 250                 255
Leu Leu Asn Trp Ala Tyr Ser Ser Gly Ala Lys Val Phe Asp Phe Pro
            260                 265                 270
Leu Tyr Tyr Lys Met Asp Glu Ala Phe Asp Asn Thr Asn Ile Pro Ala
        275                 280                 285
Leu Val Asp Ala Leu Gln Asn Gly Gly Thr Val Val Ser Arg Asp Pro
    290                 295                 300
Phe Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Asp Ile Ile Trp
305                 310                 315                 320
Asn Lys Tyr Pro Ala Tyr Ala Phe Ile Leu Thr Tyr Glu Gly Gln Pro
                325                 330                 335
Val Ile Phe Tyr Arg Asp Tyr Glu Glu Trp Leu Asn Lys Asp Lys Leu
            340                 345                 350
Asn Asn Leu Ile Trp Ile His Asp His Leu Ala Gly Gly Ser Thr Ser
        355                 360                 365
Ile Val Tyr Tyr Asp Ser Asp Glu Leu Ile Phe Val Arg Asn Gly Asp
    370                 375                 380
Ser Lys Arg Pro Gly Leu Ile Thr Tyr Ile Asn Leu Gly Ser Ser Lys
385                 390                 395                 400
Val Gly Arg Trp Val Tyr Val Pro Lys Phe Ala Gly Ala Cys Ile His
                405                 410                 415
Glu Tyr Thr Gly Asn Leu Gly Gly Trp Val Asp Lys Tyr Val Glu Ser
            420                 425                 430
Ser Gly Trp Val Tyr Leu Glu Ala Pro Ala Tyr Asp Pro Ala Ser Gly
        435                 440                 445
Gln Tyr Gly Tyr Thr Val Trp Ser Tyr Cys Gly Val Gly
    450                 455                 460
```

<210> 314
<211> 460
<212> PRT
<213> Pyrococcus furiosus

<400> 314

```
Val Asn Ile Lys Lys Leu Thr Pro Leu Leu Thr Leu Leu Leu Phe Phe
 1               5                  10                  15
Ile Val Leu Ala Ser Pro Val Ser Ala Ala Lys Tyr Leu Glu Leu Glu
            20                  25                  30
Glu Gly Gly Val Ile Met Gln Ala Phe Tyr Trp Asp Val Pro Gly Gly
        35                  40                  45
Gly Ile Trp Trp Asp His Ile Arg Ser Lys Ile Pro Glu Trp Tyr Glu
    50                  55                  60
Ala Gly Ile Ser Ala Ile Trp Leu Pro Pro Pro Ser Lys Gly Met Ser
65                  70                  75                  80
Gly Gly Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Tyr Phe Asp Leu Gly
                85                  90                  95
Glu Tyr Tyr Gln Lys Gly Thr Val Glu Thr Arg Phe Gly Ser Lys Glu
            100                 105                 110
Glu Leu Val Arg Leu Ile Gln Thr Ala His Ala Tyr Gly Ile Lys Val
        115                 120                 125
Ile Ala Asp Val Val Ile Asn His Arg Ala Gly Gly Asp Leu Glu Trp
    130                 135                 140
Asn Pro Phe Val Gly Asp Tyr Thr Trp Thr Asp Phe Ser Lys Val Ala
145                 150                 155                 160
Ser Gly Lys Tyr Thr Ala Asn Tyr Leu Asp Phe His Pro Asn Glu Leu
                165                 170                 175
His Cys Cys Asp Glu Gly Thr Phe Gly Gly Phe Pro Asp Ile Cys His
            180                 185                 190
His Lys Glu Trp Asp Gln Tyr Trp Leu Trp Lys Ser Asn Glu Ser Tyr
        195                 200                 205
Ala Ala Tyr Leu Arg Ser Ile Gly Phe Asp Gly Trp Arg Phe Asp Tyr
    210                 215                 220
Val Lys Gly Tyr Gly Ala Trp Val Val Arg Asp Trp Leu Asn Trp Trp
225                 230                 235                 240
Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn Val Asp Ala Leu
                245                 250                 255
Leu Ser Trp Ala Tyr Glu Ser Gly Ala Lys Val Phe Asp Phe Pro Leu
            260                 265                 270
Tyr Tyr Lys Met Asp Glu Ala Phe Asp Asn Asn Asn Ile Pro Ala Leu
        275                 280                 285
Val Tyr Ala Leu Gln Asn Gly Gln Thr Val Val Ser Arg Asp Pro Phe
    290                 295                 300
Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Asp Ile Ile Trp Asn
305                 310                 315                 320
Lys Tyr Pro Ala Tyr Ala Phe Ile Leu Thr Tyr Glu Gly Gln Pro Val
                325                 330                 335
Ile Phe Tyr Arg Asp Phe Glu Glu Trp Leu Asn Lys Asp Lys Leu Ile
            340                 345                 350
Asn Leu Ile Trp Ile His Asp His Leu Ala Gly Gly Ser Thr Thr Ile
        355                 360                 365
Val Tyr Tyr Asp Asn Asp Glu Leu Ile Phe Val Arg Asn Gly Asp Ser
    370                 375                 380
Arg Arg Pro Gly Leu Ile Thr Tyr Ile Asn Leu Ser Pro Asn Trp Val
385                 390                 395                 400
Gly Arg Trp Val Tyr Val Pro Lys Phe Ala Gly Ala Cys Ile His Glu
                405                 410                 415
Tyr Thr Gly Asn Leu Gly Gly Trp Val Asp Lys Arg Val Asp Ser Ser
            420                 425                 430
Gly Trp Val Tyr Leu Glu Ala Pro Pro His Asp Pro Ala Asn Gly Tyr
        435                 440                 445
Tyr Gly Tyr Ser Val Trp Ser Tyr Cys Gly Val Gly
    450                 455                 460

<210> 315
<211> 484
<212> PRT
<213> Thermococcus sp.

<400> 315
Ser Glu Ser Gln Cys Thr Ala Thr Cys Thr Trp Arg Val Val Tyr Met
 1               5                  10                  15
Ser Ala Lys Lys Leu Leu Ala Leu Leu Phe Val Leu Ala Val Leu Val
```

```
            20                  25                  30
Gly Val Ala Val Ile Pro Ala Arg Val Gly Ile Ala Pro Val Ser Ala
        35                  40                  45
Gly Ala Thr Ser Arg Pro Ser Leu Glu Glu Gly Gly Val Ile Met Gln
    50                  55                  60
Ala Phe Tyr Trp Asp Val Pro Ala Gly Gly Ile Trp Trp Asp Thr Ile
65                  70                  75                  80
Arg Ser Lys Ile Pro Asp Trp Ala Ser Ala Gly Ile Ser Ala Ile Trp
                85                  90                  95
Ile Pro Pro Ala Ser Lys Gly Met Ser Gly Ala Tyr Ser Met Gly Tyr
            100                 105                 110
Asp Pro Tyr Asp Phe Phe Asp Leu Gly Glu Tyr Tyr Gln Lys Gly Thr
        115                 120                 125
Val Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Ile Asn Met Ile Asn
    130                 135                 140
Thr Ala His Ser Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn
145                 150                 155                 160
His Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Thr Asn Ser Tyr
                165                 170                 175
Thr Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn
            180                 185                 190
Tyr Leu Asp Phe His Pro Asn Glu Val Lys Cys Cys Asp Glu Gly Thr
        195                 200                 205
Phe Gly Gly Phe Pro Asp Ile Ala His Glu Lys Ser Trp Asp Gln Tyr
    210                 215                 220
Trp Leu Trp Ala Ser Gln Lys Ser Tyr Ala Ala Tyr Leu Arg Ser Ile
225                 230                 235                 240
Gly Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp
                245                 250                 255
Val Val Lys Asp Trp Leu Lys Trp Trp Ala Leu Ala Val Gly Glu Tyr
            260                 265                 270
Trp Asp Thr Asn Val Asp Ala Leu Leu Asn Trp Ala Tyr Ser Ser Gly
        275                 280                 285
Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Glu Ala Phe
    290                 295                 300
Asp Asn Lys Asn Ile Pro Ala Leu Val Ser Ala Leu Gln Asn Gly Gln
305                 310                 315                 320
Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn
                325                 330                 335
His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile
            340                 345                 350
Leu Thr Tyr Glu Gly Gln Pro Val Ile Phe Tyr Arg Asp Tyr Glu Glu
        355                 360                 365
Trp Leu Asn Lys Asp Arg Leu Lys Asn Leu Ile Trp Ile His Asn Asn
    370                 375                 380
Leu Ala Gly Gly Ser Thr Ser Ile Val Tyr Tyr Asp Asn Asp Glu Leu
385                 390                 395                 400
Ile Phe Val Arg Asn Gly Tyr Gly Asn Lys Pro Gly Leu Ile Thr Tyr
                405                 410                 415
Ile Asn Leu Gly Ser Ser Lys Val Gly Arg Trp Val Tyr Val Pro Lys
            420                 425                 430
Phe Ala Gly Ser Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp
        435                 440                 445
Val Asp Lys Tyr Val Gly Ser Asn Gly Trp Val Tyr Leu Glu Ala Pro
    450                 455                 460
Ala His Asp Pro Ala Lys Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr
465                 470                 475                 480
Cys Gly Val Gly
```

<210> 316
<211> 457
<212> PRT
<213> Thermococcus hydrothermalis

<400> 316

```
Met Ala Arg Lys Val Leu Val Ala Leu Leu Val Phe Leu Val Val Leu
 1               5                  10                  15
```

```
Ser Val Ser Ala Val Pro Ala Lys Ala Glu Thr Leu Glu Asn Gly Gly
            20                  25                  30
Val Ile Met Gln Ala Phe Tyr Trp Asp Val Pro Gly Gly Gly Ile Trp
        35                  40                  45
Trp Asp Thr Ile Ala Gln Lys Ile Pro Asp Trp Ala Ser Ala Gly Ile
    50                  55                  60
Ser Ala Ile Trp Ile Pro Pro Ala Ser Lys Gly Met Ser Gly Gly Tyr
65                  70                  75                  80
Ser Met Gly Tyr Asp Pro Tyr Asp Phe Phe Asp Leu Gly Glu Tyr Tyr
                85                  90                  95
Gln Lys Gly Ser Val Glu Thr Arg Phe Gly Ser Lys Glu Glu Leu Val
            100                 105                 110
Asn Met Ile Asn Thr Ala His Ala His Asn Met Lys Val Ile Ala Asp
        115                 120                 125
Ile Val Ile Asn His Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe
    130                 135                 140
Thr Asn Ser Tyr Thr Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys
145                 150                 155                 160
Tyr Thr Ala Asn Tyr Leu Asp Phe His Pro Asn Glu Leu His Ala Gly
                165                 170                 175
Asp Ser Gly Thr Phe Gly Gly Tyr Pro Asp Ile Cys His Asp Lys Ser
            180                 185                 190
Trp Asp Gln His Trp Leu Trp Ala Ser Asn Glu Ser Tyr Ala Ala Tyr
        195                 200                 205
Leu Arg Ser Ile Gly Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly
    210                 215                 220
Tyr Ala Pro Trp Val Val Lys Asn Trp Leu Asn Arg Trp Gly Gly Trp
225                 230                 235                 240
Ala Val Gly Glu Tyr Trp Asp Thr Asn Val Asp Ala Leu Leu Ser Trp
                245                 250                 255
Ala Tyr Asp Ser Gly Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys
            260                 265                 270
Met Asp Glu Ala Phe Asp Asn Asn Asn Ile Pro Ala Leu Val Asp Ala
        275                 280                 285
Leu Lys Asn Gly Gly Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val
    290                 295                 300
Thr Phe Val Ala Asn His Asp Thr Asn Ile Ile Trp Asn Lys Tyr Pro
305                 310                 315                 320
Ala Tyr Ala Phe Ile Leu Thr Tyr Glu Gly Gln Pro Ala Ile Phe Tyr
                325                 330                 335
Arg Asp Tyr Glu Glu Trp Leu Asn Lys Asp Arg Leu Arg Asn Leu Ile
            340                 345                 350
Trp Ile His Asp His Leu Ala Gly Gly Ser Thr Asp Ile Ile Tyr Tyr
        355                 360                 365
Asp Ser Asp Glu Leu Ile Phe Val Arg Asn Gly Tyr Gly Asp Lys Pro
    370                 375                 380
Gly Leu Ile Thr Tyr Ile Asn Leu Gly Ser Ser Lys Ala Gly Arg Trp
385                 390                 395                 400
Val Tyr Val Pro Lys Phe Ala Gly Ser Cys Ile His Glu Tyr Thr Gly
                405                 410                 415
Asn Leu Gly Gly Trp Ile Asp Lys Trp Val Asp Ser Ser Gly Arg Val
            420                 425                 430
Tyr Leu Glu Ala Pro Ala His Asp Pro Ala Asn Gly Gln Tyr Gly Tyr
        435                 440                 445
Ser Val Trp Ser Tyr Cys Gly Val Gly
    450                 455

<210> 317
<211> 340
<212> PRT
<213> Artificial Sequence

<220>
<223> Consensus sequence

<400> 317
Ser Ala Leu Glu Gly Gly Val Ile Met Gln Ala Phe Tyr Trp Asp Val
 1               5                  10                  15
```

```
Pro Gly Gly Ile Trp Trp Asp Ile Lys Ile Pro Trp Ala Gly Ile Ser
            20                  25                  30
Ala Ile Trp Pro Pro Ser Lys Gly Met Gly Tyr Ser Met Gly Tyr Asp
        35                  40                  45
Pro Tyr Asp Phe Asp Leu Gly Glu Tyr Gln Lys Gly Glu Thr Arg Phe
    50                  55                  60
Gly Ser Lys Glu Leu Ile Thr Ala His Lys Val Ile Ala Asp Val Ile
65                  70                  75                  80
Asn His Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Tyr Thr Trp
                85                  90                  95
Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr Leu
            100                 105                 110
Asp Phe His Pro Asn Glu Asp Gly Thr Phe Gly Gly Pro Asp Ile His
        115                 120                 125
Lys Trp Asp Gln Trp Leu Trp Ser Ser Tyr Ala Ala Tyr Leu Arg Ser
    130                 135                 140
Ile Gly Asp Trp Arg Phe Asp Tyr Val Lys Gly Tyr Trp Val Val Trp
145                 150                 155                 160
Leu Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn Val Asp Ala Leu Leu
                165                 170                 175
Trp Ala Tyr Ser Gly Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys
            180                 185                 190
Met Asp Glu Ala Phe Asp Asn Asn Ile Pro Ala Leu Val Ala Leu Asn
        195                 200                 205
Gly Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala
    210                 215                 220
Asn His Asp Thr Ile Ile Trp Asn Lys Tyr Ala Tyr Ala Phe Ile Leu
225                 230                 235                 240
Thr Tyr Glu Gly Gln Pro Ile Phe Tyr Arg Asp Glu Glu Trp Leu Asn
                245                 250                 255
Lys Asp Leu Asn Leu Ile Trp Ile His Leu Ala Gly Gly Ser Thr Ile
            260                 265                 270
Tyr Tyr Asp Asp Glu Ile Phe Val Arg Asn Gly Pro Gly Leu Ile Thr
        275                 280                 285
Tyr Ile Asn Leu Gly Arg Trp Val Tyr Val Pro Lys Phe Ala Gly Cys
    290                 295                 300
Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp Asp Lys Val Ser Gly
305                 310                 315                 320
Val Tyr Glu Ala Pro Asp Pro Ala Gly Tyr Gly Tyr Val Trp Ser Tyr
                325                 330                 335
Cys Gly Val Gly
            340

<210> 318
<211> 463
<212> PRT
<213> Artificial Sequence

<220>
<223> Clone

<400> 318
Met Arg Arg Ser Ala Arg Val Leu Val Leu Ile Ile Ala Phe Phe Leu
 1               5                  10                  15
Leu Ala Gly Ile Tyr Tyr Pro Ser Thr Ser Ala Ala Lys Tyr Ser Glu
            20                  25                  30
Leu Glu Gln Gly Gly Val Ile Met Gln Ala Phe Tyr Trp Asp Val Pro
        35                  40                  45
Glu Gly Gly Ile Trp Trp Asp Thr Ile Arg Gln Lys Ile Pro Glu Trp
    50                  55                  60
Tyr Asp Ala Gly Ile Ser Ala Ile Trp Ile Pro Pro Ala Ser Lys Gly
65                  70                  75                  80
Met Gly Gly Ala Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Tyr Phe Asp
                85                  90                  95
Leu Gly Glu Phe Tyr Gln Lys Gly Thr Val Glu Thr Arg Phe Gly Ser
            100                 105                 110
Lys Glu Glu Leu Val Asn Met Ile Ser Thr Ala His Gln Tyr Gly Ile
        115                 120                 125
```

```
Lys Val Ile Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly Asp Leu
    130                 135                 140
Glu Trp Asn Pro Tyr Val Gly Asp Tyr Thr Trp Thr Asp Phe Ser Lys
145                 150                 155                 160
Val Ala Ser Gly Lys Tyr Lys Ala His Tyr Met Asp Phe His Pro Asn
                165                 170                 175
Asn Tyr Ser Thr Ser Asp Glu Gly Thr Phe Gly Gly Phe Pro Asp Ile
            180                 185                 190
Asp His Leu Val Pro Phe Asn Gln Tyr Trp Leu Trp Ala Ser Asn Glu
        195                 200                 205
Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly Ile Asp Ala Trp Arg Phe
    210                 215                 220
Asp Tyr Val Lys Gly Tyr Gly Ala Trp Val Val Lys Asp Trp Leu Ser
225                 230                 235                 240
Gln Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn Val Asp
                245                 250                 255
Ala Leu Leu Asn Trp Ala Tyr Ser Ser Gly Ala Lys Val Phe Asp Phe
            260                 265                 270
Pro Leu Tyr Tyr Lys Met Asp Glu Ala Phe Asp Asn Lys Asn Ile Pro
        275                 280                 285
Ala Leu Val Tyr Ala Ile Gln Asn Gly Glu Thr Val Val Ser Arg Asp
    290                 295                 300
Pro Phe Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Asn Ile Ile
305                 310                 315                 320
Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile Leu Thr Tyr Glu Gly Gln
                325                 330                 335
Pro Val Ile Phe Tyr Arg Asp Tyr Glu Glu Trp Leu Asn Lys Asp Lys
            340                 345                 350
Leu Asn Asn Leu Ile Trp Ile His Glu His Leu Ala Gly Gly Ser Thr
        355                 360                 365
Lys Ile Leu Tyr Tyr Asp Asp Asp Glu Leu Ile Phe Met Arg Glu Gly
    370                 375                 380
Tyr Gly Asp Arg Pro Gly Leu Ile Thr Tyr Ile Asn Leu Gly Ser Asp
385                 390                 395                 400
Trp Ala Glu Arg Trp Val Asn Val Gly Ser Lys Phe Ala Gly Tyr Thr
                405                 410                 415
Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp Val Asp Arg Tyr Val
            420                 425                 430
Gln Tyr Asp Gly Trp Val Lys Leu Thr Ala Pro Pro His Asp Pro Ala
        435                 440                 445
Asn Gly Tyr Tyr Gly Tyr Ser Val Trp Ser Tyr Ala Gly Val Gly
    450                 455                 460

<210> 319
<211> 306
<212> PRT
<213> Artificial Sequence

<220>
<223> Consensus sequence

<400> 319
Gly Gly Ile Met Gln Ala Phe Tyr Trp Asp Val Pro Gly Gly Ile Trp
 1                  5                  10                  15
Trp Asp Ile Lys Ile Pro Trp Ala Gly Ile Ser Ala Ile Trp Pro Pro
            20                  25                  30
Ser Lys Gly Met Gly Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Phe Asp
        35                  40                  45
Leu Gly Glu Gln Lys Gly Glu Thr Arg Phe Gly Ser Lys Glu Leu Ile
    50                  55                  60
Thr Ala His Lys Val Ile Ala Asp Val Ile Asn His Arg Ala Gly Gly
65                  70                  75                  80
Leu Glu Trp Asn Pro Tyr Thr Trp Thr Asp Phe Ser Lys Val Ala Ser
                85                  90                  95
Gly Lys Tyr Ala Tyr Asp Phe His Pro Asn Asp Gly Thr Phe Gly Gly
            100                 105                 110
Pro Asp Ile His Gln Trp Leu Trp Ser Ser Tyr Ala Ala Tyr Leu Arg
        115                 120                 125
```

```
Ser Ile Gly Asp Trp Phe Asp Tyr Val Lys Gly Tyr Trp Val Val Trp
    130                 135                 140
Leu Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn Val Asp Ala Leu Trp
145                 150                 155                 160
Ala Tyr Ser Ala Lys Val Phe Asp Phe Leu Tyr Tyr Lys Met Asp Ala
                165                 170                 175
Phe Asp Asn Asn Ile Pro Ala Leu Val Ala Gly Thr Val Val Ser Arg
            180                 185                 190
Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Ile Ile
        195                 200                 205
Trp Asn Lys Tyr Ala Tyr Ala Phe Ile Leu Thr Tyr Glu Gly Gln Pro
    210                 215                 220
Ile Phe Tyr Arg Asp Glu Glu Trp Leu Asn Lys Asp Leu Asn Leu Ile
225                 230                 235                 240
Trp Ile His Leu Ala Gly Gly Ser Thr Ile Tyr Tyr Asp Asp Glu Ile
                245                 250                 255
Phe Arg Gly Pro Gly Leu Ile Thr Tyr Ile Asn Leu Arg Trp Val Val
            260                 265                 270
Lys Phe Ala Gly Ile His Glu Tyr Thr Gly Leu Gly Gly Trp Asp Val
        275                 280                 285
Gly Val Ala Pro Asp Pro Ala Gly Tyr Gly Tyr Val Trp Ser Tyr Gly
    290                 295                 300
Val Gly
305
```

<210> 320
<211> 1392
<212> DNA
<213> Artificial Sequence

<220>
<223> Clone

<400> 320

```
atgaggagat ccgcaagggt tttggttctg attatagcgt ttttcctcct ggcggggatt      60
tactacccct ccacgagtgc cgcgaagtac tccgagctgg agcagggcgg agtcataatg     120
caggccttct actgggacgt tccggaggga ggaatctggt gggacacaat acggcagaag     180
atccctgaat ggtacgatgc aggcatatcc gccatctgga tacccccggc gagcaagggc     240
atgggcgggg cctactcgat gggctacgac ccctacgatt acttcgatct gggcgagttt     300
taccagaagg gaaccgttga gacccgcttc ggctccaagg aagagctcgt caacatgatc     360
tccacggccc accagtacgg catcaaggtt atagcggaca tagtgataaa ccaccgcgca     420
ggtggagacc tcgaatggaa cccatacgtc ggcgactata cctggacgga cttttctaag     480
gtcgcctccg ggaaatacaa ggcccactac atggacttcc atccaaacaa ctacagcacc     540
tcagacgagg gaaccttcgg tggcttccca gacattgatc acctcgtgcc cttcaaccag     600
tactggctgt gggcgagcaa cgagagctac gccgcctacc tcaggagcat agggatcgat     660
gcgtggcgct ttgactacgt taagggctac ggcgcgtggg tcgtcaagga ctggctgagt     720
cagtggggcg gctgggccgt cggcgagtac tgggacacca acgtcgatgc gctcctcaac     780
tgggcctaca gcagcggcgc caaggtcttc gacttcccgc tctactacaa gatggacgag     840
gcctttgaca acaagaacat tcccgccctc gtttacgcca tccagaacgg tgaaaccgtc     900
gtcagcaggg atcccttcaa ggccgttacc ttcgtggcta accacgatac gaacataatc     960
tggaacaagt accctgccta tgccttcatc ctgacctacg aaggtcagcc cgtcatcttc    1020
taccgcgact acgaggagtg gctcaacaag gacaaactca acaacctcat atggattcac    1080
gagcacctgg caggggggaag caccaagatc ctctactacg acgacgatga gctcatcttc    1140
atgagggaag gctacggcga caggcccggg cttataacct acatcaacct cggtagcgac    1200
tgggcggaga gatgggtgaa cgttggctca aagttcgcgg gctatacaat ccacgaatac    1260
accggaaacc tcggcggctg ggtcgacagg tacgtccagt acgacggctg ggtcaagctt    1320
accgctccgc cacacgatcc ggcaaacggc tattacggct actcggtctg gagctacgcc    1380
ggagttggat ga                                                        1392
```

<210> 321
<211> 846
<212> DNA
<213> Artificial Sequence

<220>
<223> Consensus sequence

<400> 321

ggcggtataa tgcaggcttc tatgggagtc cggaggaatt ggtgggacac ataagatccg 60
atgggcggat tcgcattgga tccccgcaga agggatggcg ggtatcatgg gctacgaccc 120
tacgatttga ctgggagtac agaagggatg aacgttggtc aaagacttaa atgatcacgc 180
caacataagg tatgcgaatg tataaccacc gcgcgggggc tgatggaacc tcctaacctg 240
gacgatttca agtgctcgga ataagccact actgacttcc accaacagag gacttggggt 300
ccgaatcacc taccagactg gcttgggcag cgaagctacg cgctactgga gcatggtgag 360
ctgggttgac tacgtaaggg ctagctgggt gtaactggct gtgggggggct gggcgtggga 420
gtactgggac acaacgtgag ctctactggg ctacagcggc aagtcttgac ttccctctac 480
tacaagatgg aggccttgaa aaaacatccg ctgtgctagg acgtgtccgg accttaagcg 540
tacttgtgca accacgaaca ataattggaa caagtacgct agcttcatcc tacctagagg 600
cagccatttt accgcgacta cgaggagtgg ctcaacaagg aataaacctc attggatcag 660
aacctgcggg gagacattta ctacgacgag atattttaga ggctaggcac cggctataac 720
taatcaacct gcggagtggg tagtcaagtt cgcggatcac gataacggac tcggggtggt 780
gacagtgttc gggggtatgc cccacgaccg caaggatagg ctactcgttg gagctacggg 840
tggtga 846

<210> 322
<211> 1392
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 322
atgaggagat ccgcaagggt tttggttctg attatagcgt ttttcctcct ggcggggatt 60
tactacccct ccacgagtgc cgcgaagtac tccgagctgg agcagggcgg agtcataatg 120
caggccttct actgggacgt tccggaggga ggaatctggt gggacacaat acggcagaag 180
atccctgaat ggtacgatgc aggcatatcc gccatctgga tacccccggc gagcaagggc 240
atgggcgggg cctactcgat gggctacgac ccctacgatt acttcgatct gggcgagttt 300
taccagaagg gaaccgttga gacccgcttc ggctccaagg aagagctcgt caacatgatc 360
tccacggccc accagtacgg catcaaggtt atagcggaca tagtgataaa ccaccgcgca 420
ggtggagacc tcgaatggaa cccatacgtc ggcgactata cctggacgga cttttctaag 480
gtcgcctccg ggaaatacaa ggcccactac atggacttcc atccaaacaa ctacagcacc 540
tcagacgagg gaaccttcgg tggcttccca gacattgatc acctcgtgcc cttcaaccag 600
tactggctgt gggcgagcaa cgagagctac gccgcctacc tcaggagcat agggatcgat 660
gcgtggcgct ttgactacgt taagggctac ggcgcgtggg tcgtcaagga ctggctgagt 720
cagtggggcg gctgggccgt cggcgagtac tgggacacca acgtcgatgc gctcctcaac 780
tgggcctaca gcagcggcgc caaggtcttc gacttcccgc tctactacaa gatggacgag 840
gcctttgaca acaagaacat tcccgccctc gtttacgcca tccagaacgg tgaaaccgtc 900
gtcagcaggg atcccttcaa ggccgttacc ttcgtggcta accacgatac gaacataatc 960
tggaacaagt accctgccta tgccttcatc ctgacctacg aaggtcagcc cgtcatcttc 1020
taccgcgact acgaggagtg gctcaacaag gacaaactca acaacctcat atggattcac 1080
gagcacctgg caggggaag caccaagatc ctctactacg acgacgatga gctcatcttc 1140
atgagggaag gctacggcga caggcccggg cttataacct acatcaacct cggtagcgac 1200
tgggcggaga gatgggtgaa cgttggctca aagttcgcgg gctatacaat ccacgaatac 1260
accggaaacc tcggcggctg ggtcgacagg tacgtccagt acgacggctg ggtcaagctt 1320
accgctccgc cacacgatcc ggcaaacggc tattacggct actcggtctg gagctacgcc 1380
ggagttggat ga 1392

<210> 323
<211> 463
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(27)

<400> 323
Met Arg Arg Ser Ala Arg Val Leu Val Leu Ile Ile Ala Phe Phe Leu
 1   5   10   15
Leu Ala Gly Ile Tyr Tyr Pro Ser Thr Ser Ala Ala Lys Tyr Ser Glu
            20   25   30
Leu Glu Gln Gly Gly Val Ile Met Gln Ala Phe Tyr Trp Asp Val Pro
        35   40   45

```
Glu Gly Gly Ile Trp Trp Asp Thr Ile Arg Gln Lys Ile Pro Glu Trp
    50                  55                  60
Tyr Asp Ala Gly Ile Ser Ala Ile Trp Ile Pro Pro Ala Ser Lys Gly
65                  70                  75                  80
Met Gly Gly Ala Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Tyr Phe Asp
                85                  90                  95
Leu Gly Glu Phe Tyr Gln Lys Gly Thr Val Glu Thr Arg Phe Gly Ser
            100                 105                 110
Lys Glu Glu Leu Val Asn Met Ile Ser Thr Ala His Gln Tyr Gly Ile
        115                 120                 125
Lys Val Ile Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly Asp Leu
    130                 135                 140
Glu Trp Asn Pro Tyr Val Gly Asp Tyr Thr Trp Thr Asp Phe Ser Lys
145                 150                 155                 160
Val Ala Ser Gly Lys Tyr Lys Ala His Tyr Met Asp Phe His Pro Asn
                165                 170                 175
Asn Tyr Ser Thr Ser Asp Glu Gly Thr Phe Gly Gly Phe Pro Asp Ile
            180                 185                 190
Asp His Leu Val Pro Phe Asn Gln Tyr Trp Leu Trp Ala Ser Asn Glu
        195                 200                 205
Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly Ile Asp Ala Trp Arg Phe
    210                 215                 220
Asp Tyr Val Lys Gly Tyr Gly Ala Trp Val Val Lys Asp Trp Leu Ser
225                 230                 235                 240
Gln Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn Val Asp
                245                 250                 255
Ala Leu Leu Asn Trp Ala Tyr Ser Ser Gly Ala Lys Val Phe Asp Phe
            260                 265                 270
Pro Leu Tyr Tyr Lys Met Asp Glu Ala Phe Asp Asn Lys Asn Ile Pro
        275                 280                 285
Ala Leu Val Tyr Ala Ile Gln Asn Gly Glu Thr Val Val Ser Arg Asp
    290                 295                 300
Pro Phe Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Asn Ile Ile
305                 310                 315                 320
Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile Leu Thr Tyr Glu Gly Gln
                325                 330                 335
Pro Val Ile Phe Tyr Arg Asp Tyr Glu Glu Trp Leu Asn Lys Asp Lys
            340                 345                 350
Leu Asn Asn Leu Ile Trp Ile His Glu His Leu Ala Gly Gly Ser Thr
        355                 360                 365
Lys Ile Leu Tyr Tyr Asp Asp Asp Glu Leu Ile Phe Met Arg Glu Gly
    370                 375                 380
Tyr Gly Asp Arg Pro Gly Leu Ile Thr Tyr Ile Asn Leu Gly Ser Asp
385                 390                 395                 400
Trp Ala Glu Arg Trp Val Asn Val Gly Ser Lys Phe Ala Gly Tyr Thr
                405                 410                 415
Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp Val Asp Arg Tyr Val
            420                 425                 430
Gln Tyr Asp Gly Trp Val Lys Leu Thr Ala Pro Pro His Asp Pro Ala
        435                 440                 445
Asn Gly Tyr Tyr Gly Tyr Ser Val Trp Ser Tyr Ala Gly Val Gly
    450                 455                 460
```

<210> 324
<211> 1296
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 324
atgaccagca gcctgaaaac caccagcata cccttcatct acaacatttt cccgtcgctc 60
ctcggcgagg tcgacaagtg gattccgcat gtcgagcgtg cggtggaaat gggcttcaac 120
tggatattca tcaatcccat tcatttgacc ggcgccagcg gatcgctgta cgccatccgc 180
gactatttcc agttcaaccc ggcgtttttc ccctacgact cgttcgacgc gcagaaaaac 240
cgtctgtcgg cgtttgtctc gcagtgccgc gagatggggg cggaggtgat gatcgatctg 300
gtcatcaacc ataccgccta cgacaacccc ctgacccgcg atcaccgcaa ctggtacaag 360

```
ctgaacgccg acggttcgat aaaaaatccc ggcgccaggg acgacagcgc gcccggcggg     420
tatgtggttt ggggcgactt gaacgagata gacaacgagc attccccgga ccggaacaat     480
ttgtgggaat attggtggtc ggtggtggac atgtatctgt cctgcggcgt gcgcggattc     540
cgttgcgacg cggcttatca gattccgaac gacttgtggc ggctgcttat agaacgcgcc     600
aagaacaagc attccgaatg ccgtttcttc gcggagtctc tcggctgtcc catggaacag     660
accgtcaacc tggcgaaagc gggccacgac ttcgttttca actccggcaa gtggtgggat     720
tatcgccagg actggtttat cgatcagatg cgcatggttt ccggccaggc gtcttcgatt     780
tgtttccccg agtcccacga caccgaacgt ttggcggcgg agtggggtcg ggatttggag     840
cggatcaagc agcattatct ttcagcgcg ttgatatcgt ccggcgtcct catccttctc     900
ggcttcgagt acggcttcat gaaaaagcca catgtcgtgc actcgcatcc gcacgattac     960
gagggcgcga attacgattt gaccgactac atccgggaaa tcaattccct caagcgcgcc    1020
caccgggtat tcgccgagga caataaattc gaacggcttg attccggaga cgacggcgtg    1080
ctggcgctgc ggaagacgac gctggacggc aaggagagcg gcctgcttct cttcaaccgt    1140
accggcgatc gcaaaaccct tcacctgggc gacttgatca accgcatgca tgccaacccg    1200
taccgtttca ccgcggccgc caacgtcaag ctccccgcca aggccgagcg tatcgagatc    1260
aaccggtatg gcgcggtggc gttccggttg tattga                               1296

<210> 325
<211> 431
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 325
Met Thr Ser Ser Leu Lys Thr Thr Ser Ile Pro Phe Ile Tyr Asn Ile
 1               5                  10                  15
Phe Pro Ser Leu Leu Gly Glu Val Asp Lys Trp Ile Pro His Val Glu
            20                  25                  30
Arg Ala Val Glu Met Gly Phe Asn Trp Ile Phe Ile Asn Pro Ile His
        35                  40                  45
Leu Thr Gly Ala Ser Gly Ser Leu Tyr Ala Ile Arg Asp Tyr Phe Gln
    50                  55                  60
Phe Asn Pro Ala Phe Phe Pro Tyr Asp Ser Phe Asp Ala Gln Lys Asn
65                  70                  75                  80
Arg Leu Ser Ala Phe Val Ser Gln Cys Arg Glu Met Gly Ala Glu Val
                85                  90                  95
Met Ile Asp Leu Val Ile Asn His Thr Ala Tyr Asp Asn Pro Leu Thr
            100                 105                 110
Arg Asp His Arg Asn Trp Tyr Lys Leu Asn Ala Asp Gly Ser Ile Lys
        115                 120                 125
Asn Pro Gly Ala Arg Asp Asp Ser Ala Pro Gly Gly Tyr Val Val Trp
    130                 135                 140
Gly Asp Leu Asn Glu Ile Asp Asn Glu His Ser Pro Asp Arg Asn Asn
145                 150                 155                 160
Leu Trp Glu Tyr Trp Trp Ser Val Val Asp Met Tyr Leu Ser Cys Gly
                165                 170                 175
Val Arg Gly Phe Arg Cys Asp Ala Ala Tyr Gln Ile Pro Asn Asp Leu
            180                 185                 190
Trp Arg Leu Leu Ile Glu Arg Ala Lys Asn Lys His Ser Glu Cys Arg
        195                 200                 205
Phe Phe Ala Glu Ser Leu Gly Cys Pro Met Glu Gln Thr Val Asn Leu
    210                 215                 220
Ala Lys Ala Gly His Asp Phe Val Phe Asn Ser Gly Lys Trp Trp Asp
225                 230                 235                 240
Tyr Arg Gln Asp Trp Phe Ile Asp Gln Met Arg Met Val Ser Gly Gln
                245                 250                 255
Ala Ser Ser Ile Cys Phe Pro Glu Ser His Asp Thr Glu Arg Leu Ala
            260                 265                 270
Ala Glu Trp Gly Arg Asp Leu Glu Arg Ile Lys Gln His Tyr Leu Phe
        275                 280                 285
Ser Ala Leu Ile Ser Ser Gly Val Leu Ile Leu Leu Gly Phe Glu Tyr
    290                 295                 300
Gly Phe Met Lys Lys Pro His Val Val His Ser His Pro His Asp Tyr
305                 310                 315                 320
Glu Gly Ala Asn Tyr Asp Leu Thr Asp Tyr Ile Arg Glu Ile Asn Ser
                325                 330                 335
```

Leu Lys Arg Ala His Arg Val Phe Ala Glu Asp Asn Lys Phe Glu Arg
340 345 350
Leu Asp Ser Gly Asp Asp Gly Val Leu Ala Leu Arg Lys Thr Thr Leu
355 360 365
Asp Gly Lys Glu Ser Gly Leu Leu Leu Phe Asn Arg Thr Gly Asp Arg
370 375 380
Lys Thr Leu His Leu Gly Asp Leu Ile Asn Arg Met His Ala Asn Pro
385 390 395 400
Tyr Arg Phe Thr Ala Ala Ala Asn Val Lys Leu Pro Ala Lys Ala Glu
405 410 415
Arg Ile Glu Ile Asn Arg Tyr Gly Ala Val Ala Phe Arg Leu Tyr
420 425 430

<210> 326
<211> 1359
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 326
gtggaaggac tctggctcac ccctttcttc cagtctccgt cttatcacaa atatgacgtg 60
acgaattata tggagatcga tccggaatac gggaccaaag atgatttcag gagattgctg 120
gacgaagccc acaagcgtgg cctcaaggtc atcatcgatt ttgtgatcaa ccacaccagc 180
agccagcacc cctggttcct ggaggccagg aaaggcaaag acaatcctta ccgcaattat 240
tacctgtgga tggagcccgc cgaaatcgaa aggatgggca tcgcggtgcg ggaaaaaacc 300
gccgactcct gggaaacgca gccctggcac tgggccagcc gcggcgacaa ggaaaagtat 360
tacgggatgt tctggaacgg gatgcctgat cttaatatgg acaaccgtga ggtgagagaa 420
cagatttatg ccatcggcaa atactggctc gaattcgggg ctgacggctt caggatggat 480
gctgccaaac acatttatcc cagttgggaa gccgaaaagt cgcacgactt ctgggtggag 540
ttcagagagc ggatggaggc tgtcaagcct ggcttttacc tggtgggggа ggtgtggaca 600
tcggcagaaa aggtagcccc gtttttccgg ggactgaaag ccaacttcca ttttgacctg 660
agcctggccc tgcagaaaat agcggcggga caggacagcg gatcaggact ggtgaacatg 720
ctgctgggca attacgagat attcggaaga gagaacccgg cgtttattga cgccacgatg 780
ctgaccaacc acgaccagga gaggctggga agtgcggtag gcaacgataa aaataagatg 840
aagctggctg ctaacctgct gctgaccctg cccggtaatc cctatcttta ttacggggaa 900
gaactgggga tgctcggcaa aaagcctgat gaaaacatca gggaggcgtt tctgtgggat 960
acgcgggcga atgaccgcga ccggaccaac tggaagaaac ccaattttaa taccgattcc 1020
cgggtgacac ccctgaagtt gcaacagcag gacccggact cagtttatca tcattacaag 1080
aaactgatcc gattaagaaa ggatcacccg gcattgtggc aggtatcccc gccgaatctt 1140
caggcggcac aaacggctca ggaaggtttc gtgacgttca ttcgtccgca cgaaagcggg 1200
agtttgctgg tcatacagaa cctgcagaac aacaccgccc tggtggaagt aaaagaggac 1260
atcactgaga cggtattcag cactgcaggt tttcggaata ccgaagggaa ggtcatggaa 1320
ataccggggt ttgggatggt ggtttttagg ttggggtag 1359

<210> 327
<211> 452
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 327
Met Glu Gly Leu Trp Leu Thr Pro Phe Phe Gln Ser Pro Ser Tyr His
1 5 10 15
Lys Tyr Asp Val Thr Asn Tyr Met Glu Ile Asp Pro Glu Tyr Gly Thr
20 25 30
Lys Asp Asp Phe Arg Arg Leu Leu Asp Glu Ala His Lys Arg Gly Leu
35 40 45
Lys Val Ile Ile Asp Phe Val Ile Asn His Thr Ser Ser Gln His Pro
50 55 60
Trp Phe Leu Glu Ala Arg Lys Gly Lys Asp Asn Pro Tyr Arg Asn Tyr
65 70 75 80
Tyr Leu Trp Met Glu Pro Ala Glu Ile Glu Arg Met Gly Ile Ala Val
85 90 95
Arg Glu Lys Thr Ala Asp Ser Trp Glu Thr Gln Pro Trp His Trp Ala

```
            100                 105                 110
Ser Arg Gly Asp Lys Glu Lys Tyr Tyr Gly Met Phe Trp Asn Gly Met
        115                 120                 125
Pro Asp Leu Asn Met Asp Asn Arg Glu Val Arg Glu Gln Ile Tyr Ala
    130                 135                 140
Ile Gly Lys Tyr Trp Leu Glu Phe Gly Ala Asp Gly Phe Arg Met Asp
145                 150                 155                 160
Ala Ala Lys His Ile Tyr Pro Ser Trp Glu Ala Glu Lys Ser His Asp
                165                 170                 175
Phe Trp Val Glu Phe Arg Glu Arg Met Glu Ala Val Lys Pro Gly Phe
            180                 185                 190
Tyr Leu Val Gly Glu Val Trp Thr Ser Ala Glu Lys Val Ala Pro Phe
        195                 200                 205
Phe Arg Gly Leu Lys Ala Asn Phe His Phe Asp Leu Ser Leu Ala Leu
    210                 215                 220
Gln Lys Ile Ala Ala Gly Gln Asp Ser Gly Ser Gly Leu Val Asn Met
225                 230                 235                 240
Leu Leu Gly Asn Tyr Glu Ile Phe Gly Arg Glu Asn Pro Ala Phe Ile
                245                 250                 255
Asp Ala Thr Met Leu Thr Asn His Asp Gln Glu Arg Leu Gly Ser Ala
            260                 265                 270
Val Gly Asn Asp Lys Asn Lys Met Lys Leu Ala Ala Asn Leu Leu Leu
        275                 280                 285
Thr Leu Pro Gly Asn Pro Tyr Leu Tyr Tyr Gly Glu Glu Leu Gly Met
    290                 295                 300
Leu Gly Lys Lys Pro Asp Glu Asn Ile Arg Glu Ala Phe Leu Trp Asp
305                 310                 315                 320
Thr Arg Ala Asn Asp Arg Asp Arg Thr Asn Trp Lys Lys Pro Asn Phe
                325                 330                 335
Asn Thr Asp Ser Arg Val Thr Pro Leu Lys Leu Gln Gln Gln Asp Pro
            340                 345                 350
Asp Ser Val Tyr His His Tyr Lys Lys Leu Ile Arg Leu Arg Lys Asp
        355                 360                 365
His Pro Ala Leu Trp Gln Val Ser Pro Pro Asn Leu Gln Ala Ala Gln
    370                 375                 380
Thr Ala Gln Glu Gly Phe Val Thr Phe Ile Arg Pro His Glu Ser Gly
385                 390                 395                 400
Ser Leu Leu Val Ile Gln Asn Leu Gln Asn Asn Thr Ala Leu Val Glu
                405                 410                 415
Val Lys Glu Asp Ile Thr Glu Thr Val Phe Ser Thr Ala Gly Phe Arg
            420                 425                 430
Asn Thr Glu Gly Lys Val Met Glu Ile Pro Gly Phe Gly Met Val Val
        435                 440                 445
Phe Arg Leu Gly
    450
```

<210> 328
<211> 1677
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 328
gtgcagatct cgccgccgaa cgagcacgcc tggatcacga gcggcgacgg cgcgccgtac 60
ccgtggtgga tgcgctacca gccggtcagc tacagcctcg atcggagccg cagcgggacc 120
cgcgccgagc tcgccagcat ggtgagccgg tgcgccgccg cgggcgtgca gatctacgcc 180
gacgcggtga tcaaccacat gacgggcggc tcgggcacga cgtcgagcgc cggcaaccgg 240
ccgtggggcg tgaagagcta cccgaacgtc ccgtacggcg tgaacgactt ccacgcgacg 300
tgcgcgatca ccagctacgg tgacgcaaat caggtccaga actgcgagct cgccgggctc 360
caggacctga acaccggctc ctcctacgtg cgcggcaaga tcgtcgacta cctggtcgat 420
ctctacaacc tcggcgtccg cggcttccgg atcgacgcgg ccaagcacat cagcccgacc 480
gacgtgacgg cgatcgtcgc cgccgtcgag gcccgcgtcc cgacgaagcc gttctggttc 540
ctcgaggtga tcggcgcctt cggcgaggcc gtgcagccga atcagtactt caatgtcgcc 600
gaggggcgcg cgacggtgac ggagttcgcg tacggccgcg agctgtacgg caagttcgcg 660
ggcggcggcc ggctggccga actccgcacg ttcggcgaga cgtggaacct cttgccgggc 720
gaccgcgcga tcgcgttcac cgacaaccac gacaagcagc gcggccacgg cggcggcggc 780

```
acgtacctca cgtatcacaa cggggcgacc tacgacctcg gcaacgtctt catgctggcg     840
tggccgtacg gctacccggc gctcatgtcg agctacgcgt tcaacaagac caccgattac     900
gacacgagct acgggccgcc gcacgactcg gcgaccggcg cgacgcgcgg cccgtgggac     960
ggcggcgtct ccgcccccgc gtgcttcaac cagacccgcg gcggctgggt ctgcgagcac    1020
cgcttccgcc cgatcggcaa catggtggcg ttccgcaagg cgacgatggg caactggacc    1080
gtcaccgact ggtgggacaa cggcaataac cagatcgcgt ttggccgcgg cggctcggga    1140
ttcgtggtca tcaacaagga ggcggccgcg ctgactcgca cgttcaagac gagcctcgcc    1200
gccggcaagt actgcgacgt catctcgggc gactggaacg cgggcgcggg cacctgcgcc    1260
ggcaacgtcg tgaccgtcga cgcgaccggc aacgcgagca tcaccgtcgg ggcgttcggc    1320
gccgcggcga tccacgccgg cgccaagctg ggcggcggcg gcggcggcgg caccgcctcg    1380
gtgacgttca acgaggcggc ggacaccgtc tttggccaga acatctacgt ggtcggcagc    1440
gtgccggcgc tggcgagctg gaaccccgcg agcgcggtgc cgatgacctg gatctcgggc    1500
tcggggacgc gcggcaactg gcgcgcgacg gtgacgctgc ccgcgtcgac tgcgatcgag    1560
tacaagtaca tcaagaagga cggcgcgggc accgtcatct gggagtccgg cgccaaccgc    1620
acgctgacca ccggcggcgc cggcagcacg ctgacgacga acgacagctg gaagtag       1677

<210> 329
<211> 558
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 329
Met Gln Ile Ser Pro Pro Asn Glu His Ala Trp Ile Thr Ser Gly Asp
 1               5                  10                  15
Gly Ala Pro Tyr Pro Trp Trp Met Arg Tyr Gln Pro Val Ser Tyr Ser
            20                  25                  30
Leu Asp Arg Ser Arg Ser Gly Thr Arg Ala Glu Leu Ala Ser Met Val
        35                  40                  45
Ser Arg Cys Ala Ala Ala Gly Val Gln Ile Tyr Ala Asp Ala Val Ile
    50                  55                  60
Asn His Met Thr Gly Gly Ser Gly Thr Thr Ser Ser Ala Gly Asn Arg
65                  70                  75                  80
Pro Trp Gly Val Lys Ser Tyr Pro Asn Val Pro Tyr Gly Val Asn Asp
                85                  90                  95
Phe His Ala Thr Cys Ala Ile Thr Ser Tyr Gly Asp Ala Asn Gln Val
            100                 105                 110
Gln Asn Cys Glu Leu Ala Gly Leu Gln Asp Leu Asn Thr Gly Ser Ser
        115                 120                 125
Tyr Val Arg Gly Lys Ile Val Asp Tyr Leu Val Asp Leu Tyr Asn Leu
    130                 135                 140
Gly Val Arg Gly Phe Arg Ile Asp Ala Ala Lys His Ile Ser Pro Thr
145                 150                 155                 160
Asp Val Thr Ala Ile Val Ala Ala Val Glu Ala Arg Val Pro Thr Lys
                165                 170                 175
Pro Phe Trp Phe Leu Glu Val Ile Gly Ala Phe Gly Glu Ala Val Gln
            180                 185                 190
Pro Asn Gln Tyr Phe Asn Val Ala Glu Gly Arg Ala Thr Val Thr Glu
        195                 200                 205
Phe Ala Tyr Gly Arg Glu Leu Tyr Gly Lys Phe Ala Gly Gly Gly Arg
    210                 215                 220
Leu Ala Glu Leu Arg Thr Phe Gly Glu Thr Trp Asn Leu Leu Pro Gly
225                 230                 235                 240
Asp Arg Ala Ile Ala Phe Thr Asp Asn His Asp Lys Gln Arg Gly His
                245                 250                 255
Gly Gly Gly Gly Thr Tyr Leu Thr Tyr His Asn Gly Ala Thr Tyr Asp
            260                 265                 270
Leu Gly Asn Val Phe Met Leu Ala Trp Pro Tyr Gly Tyr Pro Ala Leu
        275                 280                 285
Met Ser Ser Tyr Ala Phe Asn Lys Thr Thr Asp Tyr Asp Thr Ser Tyr
    290                 295                 300
Gly Pro Pro His Asp Ser Ala Thr Gly Ala Thr Arg Gly Pro Trp Asp
305                 310                 315                 320
Gly Gly Val Ser Ala Pro Ala Cys Phe Asn Gln Thr Arg Gly Gly Trp
                325                 330                 335
Val Cys Glu His Arg Phe Arg Pro Ile Gly Asn Met Val Ala Phe Arg
```

```
            340                 345                 350
Lys Ala Thr Met Gly Asn Trp Thr Val Thr Asp Trp Trp Asp Asn Gly
        355                 360                 365
Asn Asn Gln Ile Ala Phe Gly Arg Gly Gly Ser Gly Phe Val Val Ile
    370                 375                 380
Asn Lys Glu Ala Ala Ala Leu Thr Arg Thr Phe Lys Thr Ser Leu Ala
385                 390                 395                 400
Ala Gly Lys Tyr Cys Asp Val Ile Ser Gly Asp Trp Asn Ala Gly Ala
                405                 410                 415
Gly Thr Cys Ala Gly Asn Val Val Thr Val Asp Ala Thr Gly Asn Ala
            420                 425                 430
Ser Ile Thr Val Gly Ala Phe Gly Ala Ala Ala Ile His Ala Gly Ala
        435                 440                 445
Lys Leu Gly Gly Gly Gly Gly Gly Gly Thr Ala Ser Val Thr Phe Asn
    450                 455                 460
Glu Ala Ala Asp Thr Val Phe Gly Gln Asn Ile Tyr Val Val Gly Ser
465                 470                 475                 480
Val Pro Ala Leu Ala Ser Trp Asn Pro Ala Ser Ala Val Pro Met Thr
                485                 490                 495
Trp Ile Ser Gly Ser Gly Thr Arg Gly Asn Trp Arg Ala Thr Val Thr
            500                 505                 510
Leu Pro Ala Ser Thr Ala Ile Glu Tyr Lys Tyr Ile Lys Lys Asp Gly
        515                 520                 525
Ala Gly Thr Val Ile Trp Glu Ser Gly Ala Asn Arg Thr Leu Thr Thr
    530                 535                 540
Gly Gly Ala Gly Ser Thr Leu Thr Thr Asn Asp Ser Trp Lys
545                 550                 555
```

<210> 330
<211> 3129
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 330

```
atgataaacg aacttgaatt aactcccacc ttggaaacaa ccctaaccga aatcgtggcc      60
cggcggcgcg ttccccgggc cacctatcgt ctccagttta atgccaactt taccctggcc     120
caggccgagg ccctggtccc ttacttcgac gagctgggta ttagcgattg ttacgcctcg     180
cccctgttgg cggcctgtcc tggcagcgag cacggctatg atgtctgtga tcccagccgg     240
ctcaatccgg ccctgggcga tgagcaagcg tttgagtctt ttacctcggc cctccacgcc     300
cacaatatgg gcctcatttt agatacggtg cccaaccata tgggcattag cctgcctgag     360
aatggctggt ggttggatgt cctggaaaat ggcccggcgt caccctacgc cgcctatttt     420
gatattgact ggcatccaat caagtccgaa ctgaacaata aagtcctgct gcctattttg     480
gagaatcagt acgggcgggt gttggaaagc ggcaagttgg aactggtcta tgaaaatggg     540
gccttttta tctactatta tgacaccaag ctgccggttt ccccccgtag ctaccgcctc     600
attctaaccg agtgtctggc gaggctggat gaggaactgg ctgaggcgga tgaggatctg     660
caagagtttc gcagtatcct cacggccttg gagcacttgc cgccccgcac cgccgccgag     720
acagaggccc tggtcgaacg gagccgggaa aaagaggtgg tcaaacgccg gctggccaat     780
ctgacccaga gcaacgccac cgtgcaggcg gctatcgagg cggccgtgac caccctgaac     840
ggtcgccctg gggagccgga aagttttgac gctctacacg ccctgctcga agaccaatcg     900
taccggctgg ccttctggcg ggtcgcggcc gacgaaatta attaccgccg ctttttcgat     960
atcaacagcc tggccgccat ccgggtcgaa ttaccccagg tgtttgccga tacccacgcc    1020
ctcattttcc gctggctggc cgagggcaaa gttaccgggc tgcgtattga tcaccccgat    1080
ggcttgtgga gcccgggtgg ctacttcaaa cagttgcacg aaaattacct cctcgatcag    1140
ttggccccgg cggcaacggg cgtcgagcaa gaggaattgc atcaggccat tgtggcctgg    1200
ctggaaaccg ttgaggcgag ctggacggag gcggatcagg cagaaccggc cgaggctgct    1260
cttccctggc cgctgtacat agtggccgaa aaaattctgt ctgagacgga gccgctgccc    1320
ccggattggg ccgttgacgg caccaccggc tatgattttc tgaaccaggt caatggggta    1380
ttcatcaatc gccagagcga gcgggctttt gatcgtattt atcgccgttt taccggcatt    1440
tcggccaatt ttgccgagct ggccattcaa accaaagata tgattatgcg gctcgccctg    1500
gccagtgagg ttaacctgct ggcccatcag ttggagcaaa tcaacgaaaa gaaccgccgc    1560
taccgggatt tcaccctgag tggggtgacg gcagccctac gggaggtcct ggcctgtttg    1620
ccgatctacc ggacctacat taccccgccc catccggtca cggcccgcga cgagagctac    1680
atcgaagcgg cggtggccga ggcccggcgg cgcaacccgc gcttgagcac ctcgctgctg    1740
aactttctgc gcgatacgct cctgctgcgc aatctcgacg attttgcccc ggtcgaccgg    1800
ccccagatcg ttaatttcgt gatgaagttt cagcaagttt ccgggcctct caccgccaag    1860
```

gggctggagg acacggcttt ttacgtctat aaccggttcg tctccctcaa cgaggtcggc 1920
gggcaccccg gccagtttgg cttgagggtg gaagactttc accagcagaa tcaagagcgc 1980
ctggcccatt ggccgcacag tatgttaacc acctccactc acgataccaa gcggagtgag 2040
gatgtgcggg cgcgcctcaa tgtgttttcg gaaatgccgg aggagtggcg catggccctc 2100
tcccgctggc gccggctcaa cgaggcccat aaaatcacca tcgacggcga accggcccct 2160
gaccgcaacg acgagtatct gttttaccag accgtggtcg gggcctggcc ctttgctgtc 2220
aatgagggca tggactcggc tgcccaaagc ggcattgccc ggccggttgt ctctcggttg 2280
atgcagccgg gccggggtga gttggccggt ttccgagacc ggattctggc ttatatgcaa 2340
aaggcgacca acgaggccaa ggttcatacc agttggatca atcctaacga agcctacgac 2400
aaggccctgg aaaaatttgt gtgtgacagc cttgacaacc aacgggccaa tcgctttctg 2460
ggcgaactgc tggccttcac cgaacgggtg gcctatttcg gccagtttaa cagcctgagc 2520
caactgctgc tcaagctgac cgcccccggc gtgcccgatc tttatcaggg taccgagctg 2580
tgggatttca gcctggttga tcccgataac cgccggccgg tcgattacga ccagcggcag 2640
cgcttgctgg ccgagttaaa gcaaagtatc gccggtggcc agccggaaca aggccaaatc 2700
ctggcccaaa acttgctgga aacgagccac gatggccgga tcaaatttta tatcacagcc 2760
cgtaccctga attttcgccg ggagcagccc gagctatttg accaggggga ctaccagccg 2820
ctggcggcca gcggtgacca aagtgaccat atcatcgcct ttagccggag ctggggcgat 2880
gaaaatttta tcgtcgtcgt cccccggttg gtctttagcc tgatggcagg cgttgaacgc 2940
ccgcccctgg ccaatgtctg gggtgatacc tggctacacc tgcccgaccc ggcagccaac 3000
cgatcctacc ggcatctgtt caccggccaa acgctttcac cgcaatcaat cagtgatggc 3060
ccaggtctac cgctcgccga tgtttttgcg aattttccgg tagccttgct tgttaataaa 3120
agttcgtag 3129

<210> 331
<211> 1042
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 331
Met Ile Asn Glu Leu Glu Leu Thr Pro Thr Leu Glu Thr Thr Leu Thr
 1   5   10   15
Glu Ile Val Ala Arg Arg Arg Val Pro Arg Ala Thr Tyr Arg Leu Gln
            20   25   30
Phe Asn Ala Asn Phe Thr Leu Ala Gln Ala Glu Ala Leu Val Pro Tyr
        35   40   45
Phe Asp Glu Leu Gly Ile Ser Asp Cys Tyr Ala Ser Pro Leu Leu Ala
    50   55   60
Ala Cys Pro Gly Ser Glu His Gly Tyr Asp Val Cys Asp Pro Ser Arg
65   70   75   80
Leu Asn Pro Ala Leu Gly Asp Glu Gln Ala Phe Glu Ser Phe Thr Ser
                85   90   95
Ala Leu His Ala His Asn Met Gly Leu Ile Leu Asp Thr Val Pro Asn
            100   105   110
His Met Gly Ile Ser Leu Pro Glu Asn Gly Trp Trp Leu Asp Val Leu
        115   120   125
Glu Asn Gly Pro Ala Ser Pro Tyr Ala Ala Tyr Phe Asp Ile Asp Trp
    130   135   140
His Pro Ile Lys Ser Glu Leu Asn Asn Lys Val Leu Leu Pro Ile Leu
145   150   155   160
Glu Asn Gln Tyr Gly Arg Val Leu Glu Ser Gly Lys Leu Glu Leu Val
                165   170   175
Tyr Glu Asn Gly Ala Phe Phe Ile Tyr Tyr Tyr Asp Thr Lys Leu Pro
            180   185   190
Val Ser Pro Arg Ser Tyr Arg Leu Ile Leu Thr Glu Cys Leu Ala Arg
        195   200   205
Leu Asp Glu Glu Leu Ala Glu Ala Asp Glu Asp Leu Gln Glu Phe Arg
    210   215   220
Ser Ile Leu Thr Ala Leu Glu His Leu Pro Pro Arg Thr Ala Ala Glu
225   230   235   240
Thr Glu Ala Leu Val Glu Arg Ser Arg Glu Lys Glu Val Val Lys Arg
                245   250   255
Arg Leu Ala Asn Leu Thr Gln Ser Asn Ala Thr Val Gln Ala Ala Ile
            260   265   270
Glu Ala Ala Val Thr Thr Leu Asn Gly Arg Pro Gly Glu Pro Glu Ser
        275   280   285

```
Phe Asp Ala Leu His Ala Leu Leu Glu Asp Gln Ser Tyr Arg Leu Ala
    290                 295                 300
Phe Trp Arg Val Ala Ala Asp Glu Ile Asn Tyr Arg Arg Phe Phe Asp
305                 310                 315                 320
Ile Asn Ser Leu Ala Ala Ile Arg Val Glu Leu Pro Gln Val Phe Ala
                325                 330                 335
Asp Thr His Ala Leu Ile Phe Arg Trp Leu Ala Glu Gly Lys Val Thr
            340                 345                 350
Gly Leu Arg Ile Asp His Pro Asp Gly Leu Trp Ser Pro Gly Gly Tyr
        355                 360                 365
Phe Lys Gln Leu His Glu Asn Tyr Leu Leu Asp Gln Leu Ala Pro Ala
    370                 375                 380
Ala Thr Gly Val Glu Gln Glu Glu Leu His Gln Ala Ile Val Ala Trp
385                 390                 395                 400
Leu Glu Thr Val Glu Ala Ser Trp Thr Glu Ala Asp Gln Ala Glu Pro
                405                 410                 415
Ala Glu Ala Ala Leu Pro Trp Pro Leu Tyr Ile Val Ala Glu Lys Ile
            420                 425                 430
Leu Ser Glu Thr Glu Pro Leu Pro Pro Asp Trp Ala Val Asp Gly Thr
        435                 440                 445
Thr Gly Tyr Asp Phe Leu Asn Gln Val Asn Gly Val Phe Ile Asn Arg
    450                 455                 460
Gln Ser Glu Arg Ala Phe Asp Arg Ile Tyr Arg Arg Phe Thr Gly Ile
465                 470                 475                 480
Ser Ala Asn Phe Ala Glu Leu Ala Ile Gln Thr Lys Asp Met Ile Met
                485                 490                 495
Arg Leu Ala Leu Ala Ser Glu Val Asn Leu Leu Ala His Gln Leu Glu
            500                 505                 510
Gln Ile Asn Glu Lys Asn Arg Arg Tyr Arg Asp Phe Thr Leu Ser Gly
        515                 520                 525
Val Thr Ala Ala Leu Arg Glu Val Leu Ala Cys Leu Pro Ile Tyr Arg
    530                 535                 540
Thr Tyr Ile Thr Pro Pro His Pro Val Thr Ala Arg Asp Glu Ser Tyr
545                 550                 555                 560
Ile Glu Ala Ala Val Ala Glu Ala Arg Arg Arg Asn Pro Arg Leu Ser
                565                 570                 575
Thr Ser Leu Leu Asn Phe Leu Arg Asp Thr Leu Leu Leu Arg Asn Leu
            580                 585                 590
Asp Asp Phe Ala Pro Val Asp Arg Pro Gln Ile Val Asn Phe Val Met
        595                 600                 605
Lys Phe Gln Gln Val Ser Gly Pro Leu Thr Ala Lys Gly Leu Glu Asp
    610                 615                 620
Thr Ala Phe Tyr Val Tyr Asn Arg Phe Val Ser Leu Asn Glu Val Gly
625                 630                 635                 640
Gly His Pro Gly Gln Phe Gly Leu Arg Val Glu Asp Phe His Gln Gln
                645                 650                 655
Asn Gln Glu Arg Leu Ala His Trp Pro His Ser Met Leu Thr Thr Ser
            660                 665                 670
Thr His Asp Thr Lys Arg Ser Glu Asp Val Arg Ala Arg Leu Asn Val
        675                 680                 685
Phe Ser Glu Met Pro Glu Glu Trp Arg Met Ala Leu Ser Arg Trp Arg
    690                 695                 700
Arg Leu Asn Glu Ala His Lys Ile Thr Ile Asp Gly Glu Pro Ala Pro
705                 710                 715                 720
Asp Arg Asn Asp Glu Tyr Leu Phe Tyr Gln Thr Val Val Gly Ala Trp
                725                 730                 735
Pro Phe Ala Val Asn Glu Gly Met Asp Ser Ala Ala Gln Ser Gly Ile
            740                 745                 750
Ala Arg Pro Val Val Ser Arg Leu Met Gln Pro Gly Arg Gly Glu Leu
        755                 760                 765
Ala Gly Phe Arg Asp Arg Ile Leu Ala Tyr Met Gln Lys Ala Thr Asn
    770                 775                 780
Glu Ala Lys Val His Thr Ser Trp Ile Asn Pro Asn Glu Ala Tyr Asp
785                 790                 795                 800
Lys Ala Leu Glu Lys Phe Val Cys Asp Ser Leu Asp Asn Gln Arg Ala
                805                 810                 815
Asn Arg Phe Leu Gly Glu Leu Leu Ala Phe Thr Glu Arg Val Ala Tyr
            820                 825                 830
```

```
Phe Gly Gln Phe Asn Ser Leu Ser Gln Leu Leu Leu Lys Leu Thr Ala
        835                 840                 845
Pro Gly Val Pro Asp Leu Tyr Gln Gly Thr Glu Leu Trp Asp Phe Ser
    850                 855                 860
Leu Val Asp Pro Asp Asn Arg Arg Pro Val Asp Tyr Asp Gln Arg Gln
865                 870                 875                 880
Arg Leu Leu Ala Glu Leu Lys Gln Ser Ile Ala Gly Gly Gln Pro Glu
                885                 890                 895
Gln Gly Gln Ile Leu Ala Gln Asn Leu Leu Glu Thr Ser His Asp Gly
            900                 905                 910
Arg Ile Lys Phe Tyr Ile Thr Ala Arg Thr Leu Asn Phe Arg Arg Glu
        915                 920                 925
Gln Pro Glu Leu Phe Asp Gln Gly Asp Tyr Gln Pro Leu Ala Ala Ser
    930                 935                 940
Gly Asp Gln Ser Asp His Ile Ile Ala Phe Ser Arg Ser Trp Gly Asp
945                 950                 955                 960
Glu Asn Phe Ile Val Val Val Pro Arg Leu Val Phe Ser Leu Met Ala
                965                 970                 975
Gly Val Glu Arg Pro Pro Leu Ala Asn Val Trp Gly Asp Thr Trp Leu
            980                 985                 990
His Leu Pro Asp Pro Ala Ala Asn Arg Ser Tyr Arg His Leu Phe Thr
        995                 1000                1005
Gly Gln Thr Leu Ser Pro Gln Ser Ile Ser Asp Gly Pro Gly Leu Pro
    1010                1015                1020
Leu Ala Asp Val Phe Ala Asn Phe Pro Val Ala Leu Leu Val Asn Lys
1025                1030                1035                1040
Ser Ser
```

<210> 332
<211> 1707
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 332

```
atgtcgaaca agaaggcgtt gggggcggta ctcgcgacca ctctgctgcc ggggttggcg      60
ctggcgggca gcagtgacgt gttgatgcag ggcttccact ggaagtccac caaccaggtg     120
ggctggtaca acaccgtcaa gagcaaggcg ggcgagctcg gagcggctgg catcaaccat     180
gtctggctcc cgccgccctc caagtcggga gcgccccagg gctaccttcc cacggagctc     240
aacgtgctgg actcggcgta cggcacggag ctgcaactga aggacgccat ctccgcgctg     300
aacgcggcgg gggtgagcgc ggtggcggac atcgtggtga accatcgcaa cggctcgacg     360
gcgtggcacg acttccgcaa cccggactgg ggcacggact ccatcgtggc caatgacgag     420
tgctggagca cggcgggctc cacctgcacg agcagcatgg tgcgcggggc caacgacagc     480
ggcgagccgt acgccgccgc gcgcgacgtg gaccacggca agcagtacgt gcgcgacagc     540
ctcgtgggct ggatggggac gcggctggcg aacgtgggct tcgacggctg gcgcttcgac     600
ttcgtgaagg gcttcagcgg cgcgtacgtg ggcgagtacg tggcccgcac cgcgccgacg     660
ttctgcgtgg gcgagttctg gccgacgaac tacttcgacc tgaatgaccc ggccaactgg     720
cgcaaccaga ttatcggctg ggtggatgcc acgggcggcc ggtgcagcgc cttcgacttc     780
gtcaccaagg gcctgctcaa ccaggtgctg gccaacggcg actacgggcg gctgaagacg     840
gcggacggca agccgacggg aacgattggc gtgcggccgg cgcgctcggt caccttcgtg     900
gacaaccacg acacgggccc cagcgaggcg tgcggggatg gtcagaacca ctggccggtg     960
ccgtgcgaca aggtgatggc gggctacgcc tacatcctca cgcacccggg tgtgccgacg    1020
gtgtactgga cgcactatta caacttcggc ctgaaggcgc agctcgacaa gctcatcagc    1080
atccgcaaga cgaacggcct gcagtcggag tcggtggtca acatcgtccg ggccgagcag    1140
aacctgtacg cggccatcat cgacgccaag gtggcgatga agatcggctc ggggagctgg    1200
agcccgggca cggactggac gctggccgct tccggtacgg actacgcggt gtggaccaag    1260
ggcaccacgc ctccgccccc caccggctgc acccagccgt cgatgaacct gcgcggcacg    1320
ttcaacaact tcggcagcgc cgccatgacg tgcaagggcg gcaacacctg ggagatcacc    1380
aacgtcacct tcggtggcgc cgcgacggac cgcttcaagt tcgacgtgtt tggtgactgg    1440
acccggaact atggcgacac caacaacgat gggaccgccg agctgtcggg tggggacatc    1500
tccctcacca ccgcgggcac ctacaccctc cagttcaacg acagcacgct gcgctacacg    1560
gtcacccgca ccagcacccc accgccgccc acgggcacca gcgttcgctt cgtctgccag    1620
aacggcacca ccttcgtggg acagaacgtg cacgtcgtgg gcagcatccc cgagctgggg    1680
agctgggatg cgacgaagtc catcaag                                        1707
```

<210> 333
<211> 569
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(22)

<400> 333
```
Met Ser Asn Lys Lys Ala Leu Gly Ala Val Leu Ala Thr Thr Leu Leu
 1               5                  10                  15
Pro Gly Leu Ala Leu Ala Gly Ser Ser Asp Val Leu Met Gln Gly Phe
            20                  25                  30
His Trp Lys Ser Thr Asn Gln Val Gly Trp Tyr Asn Thr Val Lys Ser
        35                  40                  45
Lys Ala Gly Glu Leu Gly Ala Ala Gly Ile Asn His Val Trp Leu Pro
    50                  55                  60
Pro Pro Ser Lys Ser Gly Ala Pro Gln Gly Tyr Leu Pro Thr Glu Leu
65                  70                  75                  80
Asn Val Leu Asp Ser Ala Tyr Gly Thr Glu Leu Gln Leu Lys Asp Ala
                85                  90                  95
Ile Ser Ala Leu Asn Ala Ala Gly Val Ser Ala Val Ala Asp Ile Val
            100                 105                 110
Val Asn His Arg Asn Gly Ser Thr Ala Trp His Asp Phe Arg Asn Pro
        115                 120                 125
Asp Trp Gly Thr Asp Ser Ile Val Ala Asn Asp Glu Cys Trp Ser Thr
    130                 135                 140
Ala Gly Ser Thr Cys Thr Ser Ser Met Val Arg Gly Ala Asn Asp Ser
145                 150                 155                 160
Gly Glu Pro Tyr Ala Ala Ala Arg Asp Val Asp His Gly Lys Gln Tyr
                165                 170                 175
Val Arg Asp Ser Leu Val Gly Trp Met Gly Thr Arg Leu Ala Asn Val
            180                 185                 190
Gly Phe Asp Gly Trp Arg Phe Asp Phe Val Lys Gly Phe Ser Gly Ala
        195                 200                 205
Tyr Val Gly Glu Tyr Val Ala Arg Thr Ala Pro Thr Phe Cys Val Gly
    210                 215                 220
Glu Phe Trp Pro Thr Asn Tyr Phe Asp Leu Asn Asp Pro Ala Asn Trp
225                 230                 235                 240
Arg Asn Gln Ile Ile Gly Trp Val Asp Ala Thr Gly Gly Arg Cys Ser
                245                 250                 255
Ala Phe Asp Phe Val Thr Lys Gly Leu Leu Asn Gln Val Leu Ala Asn
            260                 265                 270
Gly Asp Tyr Gly Arg Leu Lys Thr Ala Asp Gly Lys Pro Thr Gly Thr
        275                 280                 285
Ile Gly Val Arg Pro Ala Arg Ser Val Thr Phe Val Asp Asn His Asp
    290                 295                 300
Thr Gly Pro Ser Glu Ala Cys Gly Asp Gly Gln Asn His Trp Pro Val
305                 310                 315                 320
Pro Cys Asp Lys Val Met Ala Gly Tyr Ala Tyr Ile Leu Thr His Pro
                325                 330                 335
Gly Val Pro Thr Val Tyr Trp Thr His Tyr Tyr Asn Phe Gly Leu Lys
            340                 345                 350
Ala Gln Leu Asp Lys Leu Ile Ser Ile Arg Lys Thr Asn Gly Leu Gln
        355                 360                 365
Ser Glu Ser Val Val Asn Ile Val Arg Ala Glu Gln Asn Leu Tyr Ala
    370                 375                 380
Ala Ile Ile Asp Ala Lys Val Ala Met Lys Ile Gly Ser Gly Ser Trp
385                 390                 395                 400
Ser Pro Gly Thr Asp Trp Thr Leu Ala Ala Ser Gly Thr Asp Tyr Ala
                405                 410                 415
Val Trp Thr Lys Gly Thr Thr Pro Pro Pro Pro Thr Gly Cys Thr Gln
            420                 425                 430
Pro Ser Met Asn Leu Arg Gly Thr Phe Asn Asn Phe Gly Ser Ala Ala
        435                 440                 445
```

```
Met Thr Cys Lys Gly Gly Asn Thr Trp Glu Ile Thr Asn Val Thr Phe
    450                 455                 460
Gly Gly Ala Ala Thr Asp Arg Phe Lys Phe Asp Val Phe Gly Asp Trp
465                 470                 475                 480
Thr Arg Asn Tyr Gly Asp Thr Asn Asn Asp Gly Thr Ala Glu Leu Ser
                485                 490                 495
Gly Gly Asp Ile Ser Leu Thr Thr Ala Gly Thr Tyr Thr Leu Gln Phe
            500                 505                 510
Asn Asp Ser Thr Leu Arg Tyr Thr Val Thr Arg Thr Ser Thr Pro Pro
        515                 520                 525
Pro Pro Thr Gly Thr Ser Val Arg Phe Val Cys Gln Asn Gly Thr Thr
    530                 535                 540
Phe Val Gly Gln Asn Val His Val Val Gly Ser Ile Pro Glu Leu Gly
545                 550                 555                 560
Ser Trp Asp Ala Thr Lys Ser Ile Lys
                565
```

<210> 334
<211> 2061
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 334

```
atggtagcag gctttgggct ctacggagct gccttgctca ccccaatggc ggcacaggcc      60
gaaacgggca gagaagtcat gctgcagggc tttaactgga attccacctc gcaagcgggt     120
gggcactata atgagctggc aaatcgtgct ggtgaaatcg ctggggctgg tatcgacatt     180
gtttggttcc cacctccatc cagagccgca gaccgggttg gatacctgcc caacgagtgg     240
tataatatga actccaatta tggaaatcgg acgactctac aggcggctat cgggaatctt     300
cgcaacaacg gtgtcaaaac cgtcgcggat attgtcgtta accaccgtgt gggcacaacg     360
aattgggcgg acttcacgaa tccttctttc ggtgataaca accgtgccat cacacgcgat     420
gatgaatggc atcagtcgtc gggtaactgg gataccggtg aggcgtatag tgccgcccgt     480
gacttggacc acacctacgg ccccgtgcaa aatgagatta aaaactggct gaattggttg     540
aagagtgaca tcggctttga tggttggcgc tatgatatgg tcaagggatt cagcgggtac     600
tatgtgggcg agtacaacac agcaactagc ccctacattt ccgttggtga gttcttcgac     660
tatgatcgcc agaaggtggt cggctggatc aacgccacca atgctcgctc tcgcgccttc     720
gattttccaa cacgcaacct cctctatgtg gctgtcaccc aaaacaacta cggagttctt     780
cgtgatggtg aaggtaaggc caatggactc atcggttggt ggccacaacg ggctatcaca     840
ttcattgaga accatgatac ggaagaagcc cgcaacggtg agtacacccc cgccttcccc     900
caatgggcta ccatgcaggg ttatgcctac attttgacgc accccggcat tccatgcgtt     960
ttctggaatg attggcgttg ggacttccgt tcggaaatca atcaactgat cgccatccgc    1020
agggcccaag gcatcaacga tggcagcagc ctcagcatcc aggttgccga tggcagccgc    1080
tacggggcga tcatcaatgg aaacacggcg gtcaagattg gtcctggaaa ctggagccca    1140
agtggtagtt ggacgctggc tgccgccgga accaattacg ccgtctggac aaacggtggc    1200
ggaacaccaa cgccaacacc cacaccaacc ccaactcagg gccccacaac ggttacttgg    1260
aatccatcca ctccgacagc cggacaaaat gtgacgatca cttatccttc cggccgctcc    1320
cttgccagtt cctccaacgt caacctctac tggggtgtca atggatggac caatgtccaa    1380
accaaagcga tgaccaagaa cagttccaat gactggacaa ccacgatcac acttccctcc    1440
aacacaacac gtttgaactt tgtgttcaat aacggctcca gttgggacaa caacagcagt    1500
caggattgga atgtcaacgt gaccgctgta acgcccacac caacaccgac accgactccc    1560
acacctacgg caactccgac accgacaccc accccgacgc ccactgccac acccacccca    1620
actccgactc cttcgatcat ttggtatcgc atcgaagccc gtcacagtgg taaagttcta    1680
gacgttgcca gtgcatcaac ttcaaacggt ggcaatgtgc atcagtggtc ttatgctggt    1740
ggacagaacc aacaatggcg tgttgtcgac gctgggaacg ggtttgttta catcctaaac    1800
cggaacagtg gaaaagcact tgaagtgggc aacttctcca ccagtaacgg cgggaacgtc    1860
cagcaatggg attatgccgg tggttccagt caacagtgga agctcattga gacgaccaac    1920
ggatatgtgc aaatccagaa tcggaacagc ggcaaagcca ttgatgtttc cgctgcttcg    1980
actaccaatg gtgccaacat ccatcagtgg acctacggcg gcggcaacaa tcagcagtgg    2040
aagctgatcc caatcaatta a                                              2061
```

<210> 335
<211> 686
<212> PRT
<213> Unknown

<220>

<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(20)

<400> 335

```
Met Val Ala Gly Phe Gly Leu Tyr Gly Ala Ala Leu Leu Thr Pro Met
 1               5                  10                  15
Ala Ala Gln Ala Glu Thr Gly Arg Glu Val Met Leu Gln Gly Phe Asn
            20                  25                  30
Trp Asn Ser Thr Ser Gln Ala Gly Gly His Tyr Asn Glu Leu Ala Asn
        35                  40                  45
Arg Ala Gly Glu Ile Ala Gly Ala Gly Ile Asp Ile Val Trp Phe Pro
    50                  55                  60
Pro Pro Ser Arg Ala Ala Asp Arg Val Gly Tyr Leu Pro Asn Glu Trp
65                  70                  75                  80
Tyr Asn Met Asn Ser Asn Tyr Gly Asn Arg Thr Thr Leu Gln Ala Ala
                85                  90                  95
Ile Gly Asn Leu Arg Asn Asn Gly Val Lys Thr Val Ala Asp Ile Val
            100                 105                 110
Val Asn His Arg Val Gly Thr Thr Asn Trp Ala Asp Phe Thr Asn Pro
        115                 120                 125
Ser Phe Gly Asp Asn Asn Arg Ala Ile Thr Arg Asp Asp Glu Trp His
    130                 135                 140
Gln Ser Ser Gly Asn Trp Asp Thr Gly Glu Ala Tyr Ser Ala Ala Arg
145                 150                 155                 160
Asp Leu Asp His Thr Tyr Gly Pro Val Gln Asn Glu Ile Lys Asn Trp
                165                 170                 175
Leu Asn Trp Leu Lys Ser Asp Ile Gly Phe Asp Gly Trp Arg Tyr Asp
            180                 185                 190
Met Val Lys Gly Phe Ser Gly Tyr Tyr Val Gly Glu Tyr Asn Thr Ala
        195                 200                 205
Thr Ser Pro Tyr Ile Ser Val Gly Glu Phe Phe Asp Tyr Asp Arg Gln
    210                 215                 220
Lys Val Val Gly Trp Ile Asn Ala Thr Asn Ala Arg Ser Arg Ala Phe
225                 230                 235                 240
Asp Phe Pro Thr Arg Asn Leu Leu Tyr Val Ala Val Thr Gln Asn Asn
                245                 250                 255
Tyr Gly Val Leu Arg Asp Gly Glu Gly Lys Ala Asn Gly Leu Ile Gly
            260                 265                 270
Trp Trp Pro Gln Arg Ala Ile Thr Phe Ile Glu Asn His Asp Thr Glu
        275                 280                 285
Glu Ala Arg Asn Gly Glu Tyr Thr Pro Ala Phe Pro Gln Trp Ala Thr
    290                 295                 300
Met Gln Gly Tyr Ala Tyr Ile Leu Thr His Pro Gly Ile Pro Cys Val
305                 310                 315                 320
Phe Trp Asn Asp Trp Arg Trp Asp Phe Arg Ser Glu Ile Asn Gln Leu
                325                 330                 335
Ile Ala Ile Arg Arg Ala Gln Gly Ile Asn Asp Gly Ser Ser Leu Ser
            340                 345                 350
Ile Gln Val Ala Asp Gly Ser Arg Tyr Gly Ala Ile Ile Asn Gly Asn
        355                 360                 365
Thr Ala Val Lys Ile Gly Pro Gly Asn Trp Ser Pro Ser Gly Ser Trp
    370                 375                 380
Thr Leu Ala Ala Ala Gly Thr Asn Tyr Ala Val Trp Thr Asn Gly Gly
385                 390                 395                 400
Gly Thr Pro Thr Pro Thr Pro Thr Pro Thr Pro Thr Gln Gly Pro Thr
                405                 410                 415
Thr Val Thr Trp Asn Pro Ser Thr Pro Thr Ala Gly Gln Asn Val Thr
            420                 425                 430
Ile Thr Tyr Pro Ser Gly Arg Ser Leu Ala Ser Ser Ser Asn Val Asn
        435                 440                 445
Leu Tyr Trp Gly Val Asn Gly Trp Thr Asn Val Gln Thr Lys Ala Met
    450                 455                 460
Thr Lys Asn Ser Ser Asn Asp Trp Thr Thr Thr Ile Thr Leu Pro Ser
465                 470                 475                 480
Asn Thr Thr Arg Leu Asn Phe Val Phe Asn Asn Gly Ser Ser Trp Asp
                485                 490                 495
```

```
Asn Asn Ser Ser Gln Asp Trp Asn Val Asn Val Thr Ala Val Thr Pro
            500                 505                 510
Thr Pro Thr Pro Thr Pro Thr Pro Thr Pro Thr Ala Thr Pro Thr Pro
        515                 520                 525
Thr Pro Thr Pro Thr Pro Thr Ala Thr Pro Thr Pro Thr Pro Thr Pro
    530                 535                 540
Ser Ile Ile Trp Tyr Arg Ile Glu Ala Arg His Ser Gly Lys Val Leu
545                 550                 555                 560
Asp Val Ala Ser Ala Ser Thr Ser Asn Gly Gly Asn Val His Gln Trp
                565                 570                 575
Ser Tyr Ala Gly Gly Gln Asn Gln Gln Trp Arg Val Val Asp Ala Gly
            580                 585                 590
Asn Gly Phe Val Tyr Ile Leu Asn Arg Asn Ser Gly Lys Ala Leu Glu
        595                 600                 605
Val Gly Asn Phe Ser Thr Ser Asn Gly Gly Asn Val Gln Gln Trp Asp
    610                 615                 620
Tyr Ala Gly Gly Ser Ser Gln Gln Trp Lys Leu Ile Glu Thr Thr Asn
625                 630                 635                 640
Gly Tyr Val Gln Ile Gln Asn Arg Asn Ser Gly Lys Ala Ile Asp Val
                645                 650                 655
Ser Ala Ala Ser Thr Thr Asn Gly Ala Asn Ile His Gln Trp Thr Tyr
            660                 665                 670
Gly Gly Gly Asn Asn Gln Gln Trp Lys Leu Ile Pro Ile Asn
        675                 680                 685

<210> 336
<211> 1731
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 336
atgaccgaac gacacctcag tcgcagggcg aaggggcgt tggcgggcct cgccgctgcc      60
gccgtcgccg ccgtcaccct cgcggctccc cccgcagcgc aggcatcgcc ccccgggaac     120
aaggacgtca ccgctgtcct gttctcgtgg aacttcgact ccatagcccg ggagtgcacc     180
gaccgcctgg gccccgccgg ctacggcttc gtccaggtct caccgcccca ggagcacatc     240
cagggctccg cgtggtggac ccagtaccag cccgtcagct accggctgga gagcaggctc     300
ggcaaccgcg cccagttccg gaacatggtc gacacctgca acagcgcggg tgtgggcgtg     360
gtcgtcgaca ccgtcatcaa ccacatgacc gcggggtccg gcaccggcac cggcggttcg     420
cagtacagca agtaccacta ccccgacgcc ccttacagcg acggggactt caacgactgc     480
cgccgcgaca tcagcaacta ccgcgaccgg tacgacgtcc agaactgcga actggtcggc     540
ctcgcggacc tgcggaccgg caaggagtac gtccgccagc gcatcgccga ctacatgaac     600
gacctgctgt cgctgggcgt cgccggcttc cgtatcgacg ccgccaagca catgcccgcc     660
gacgacctgc gggccatcag ggccaagctg aacaacccgg acgcctactg gaagcaggaa     720
gtcatccatg gcgccggcga ggccgtccag cccgaggagt acctcggcgc cggcgacgtc     780
caggagttcc gctacgcgca cgacatgaag cgcatgttcc agcaggaacg catcgcccac     840
ctgcggacct tcggcgagag ctggggcttc atgcccagct cgcagtcggc ggtcttcgtc     900
gacaaccacg acaccgagcg caacggcacc accctcagct acaaggacgg gtcggcctac     960
acgctcgcca acgtcttcat gctggcctgg ccctacggct ccccggacgt ccactccggg    1020
taccggttca gcaacttcga cgcgggccct cccggcaacg gccaggtgag ccagtgctgg    1080
cagaacggct gggagtgcca gcacgcctgg cccgagatcg agtccatggt cgccttccgc    1140
aacgccaccc gcggccagca ggtcaccaac tggtgggaca acggcaacaa cgtcatcgcc    1200
ttcggccggg gcagccaggg ctacgtggcg atcaaccggg agagcggacc ggtcacccgc    1260
accttccaga cctccctggc cgccggtgac tactgcgacg tccagtccgg gtccacggtc    1320
accgtgaacg gctcggggca gttcaccgcc accgtggccg gcggaaccgc gctcgccctg    1380
cacaccggcg cccggaactg caccggcaac ggcggcggca ccccccagga gccggtggac    1440
ggcggcgtct ccttcgccgt caacgcctcc acgcagtggg gcgagcacct ctatgtcacc    1500
ggcaaccgcc ccgaactggg caactgggac ccgcagaacg ccgtccggct cggctcggac    1560
gactacccgg tctggaaggc cgaactggcg atgcccgccg acacctcctt ccagtacaag    1620
tacctgcggc gcaccgacgg cggaaccgtg acctgggagc agggcggcaa ccgcaccgcg    1680
accgtcccgg cctccgggac cctgcggctg aacgacacct ggcgcagcta g              1731

<210> 337
<211> 576
<212> PRT
<213> Unknown
```

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(35)

<400> 337
Met Thr Glu Arg His Leu Ser Arg Arg Ala Lys Gly Ala Leu Ala Gly
 1   5   10   15
Leu Ala Ala Ala Ala Val Ala Ala Val Thr Leu Ala Ala Pro Pro Ala
            20   25   30
Ala Gln Ala Ser Pro Pro Gly Asn Lys Asp Val Thr Ala Val Leu Phe
        35   40   45
Ser Trp Asn Phe Asp Ser Ile Ala Arg Glu Cys Thr Asp Arg Leu Gly
    50   55   60
Pro Ala Gly Tyr Gly Phe Val Gln Val Ser Pro Pro Gln Glu His Ile
65   70   75   80
Gln Gly Ser Ala Trp Trp Thr Gln Tyr Gln Pro Val Ser Tyr Arg Leu
                85   90   95
Glu Ser Arg Leu Gly Asn Arg Ala Gln Phe Arg Asn Met Val Asp Thr
            100   105   110
Cys Asn Ser Ala Gly Val Gly Val Val Val Asp Thr Val Ile Asn His
        115   120   125
Met Thr Ala Gly Ser Gly Thr Gly Thr Gly Gly Ser Gln Tyr Ser Lys
    130   135   140
Tyr His Tyr Pro Asp Ala Pro Tyr Ser Asp Gly Asp Phe Asn Asp Cys
145   150   155   160
Arg Arg Asp Ile Ser Asn Tyr Arg Asp Arg Tyr Asp Val Gln Asn Cys
                165   170   175
Glu Leu Val Gly Leu Ala Asp Leu Arg Thr Gly Lys Glu Tyr Val Arg
            180   185   190
Gln Arg Ile Ala Asp Tyr Met Asn Asp Leu Leu Ser Leu Gly Val Ala
        195   200   205
Gly Phe Arg Ile Asp Ala Ala Lys His Met Pro Ala Asp Asp Leu Arg
    210   215   220
Ala Ile Arg Ala Lys Leu Asn Asn Pro Asp Ala Tyr Trp Lys Gln Glu
225   230   235   240
Val Ile His Gly Ala Gly Glu Ala Val Gln Pro Glu Glu Tyr Leu Gly
                245   250   255
Ala Gly Asp Val Gln Glu Phe Arg Tyr Ala His Asp Met Lys Arg Met
            260   265   270
Phe Gln Gln Glu Arg Ile Ala His Leu Arg Thr Phe Gly Glu Ser Trp
        275   280   285
Gly Phe Met Pro Ser Ser Gln Ser Ala Val Phe Val Asp Asn His Asp
    290   295   300
Thr Glu Arg Asn Gly Thr Thr Leu Ser Tyr Lys Asp Gly Ser Ala Tyr
305   310   315   320
Thr Leu Ala Asn Val Phe Met Leu Ala Trp Pro Tyr Gly Ser Pro Asp
                325   330   335
Val His Ser Gly Tyr Arg Phe Ser Asn Phe Asp Ala Gly Pro Pro Gly
            340   345   350
Asn Gly Gln Val Ser Gln Cys Trp Gln Asn Gly Trp Glu Cys Gln His
        355   360   365
Ala Trp Pro Glu Ile Glu Ser Met Val Ala Phe Arg Asn Ala Thr Arg
    370   375   380
Gly Gln Gln Val Thr Asn Trp Trp Asp Asn Gly Asn Asn Val Ile Ala
385   390   395   400
Phe Gly Arg Gly Ser Gln Gly Tyr Val Ala Ile Asn Arg Glu Ser Gly
                405   410   415
Pro Val Thr Arg Thr Phe Gln Thr Ser Leu Ala Ala Gly Asp Tyr Cys
            420   425   430
Asp Val Gln Ser Gly Ser Thr Val Thr Val Asn Gly Ser Gly Gln Phe
        435   440   445
Thr Ala Thr Val Ala Gly Gly Thr Ala Leu Ala Leu His Thr Gly Ala
    450   455   460
Arg Asn Cys Thr Gly Asn Gly Gly Gly Thr Pro Gln Glu Pro Val Asp
465   470   475   480

```
Gly Gly Val Ser Phe Ala Val Asn Ala Ser Thr Gln Trp Gly Glu His
                485                 490                 495
Leu Tyr Val Thr Gly Asn Arg Pro Glu Leu Gly Asn Trp Asp Pro Gln
            500                 505                 510
Asn Ala Val Arg Leu Gly Ser Asp Asp Tyr Pro Val Trp Lys Ala Glu
        515                 520                 525
Leu Ala Met Pro Ala Asp Thr Ser Phe Gln Tyr Lys Tyr Leu Arg Arg
    530                 535                 540
Thr Asp Gly Gly Thr Val Thr Trp Glu Gln Gly Gly Asn Arg Thr Ala
545                 550                 555                 560
Thr Val Pro Ala Ser Gly Thr Leu Arg Leu Asn Asp Thr Trp Arg Ser
                565                 570                 575
```

<210> 338
<211> 1704
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 338

```
atgccgcacc aatacaacta caggccgaaa atacaccgat tattttcagg tgctccaatg      60
caaacaaaca agaaaacctt cctggcgcat tgccttcgcc aaagtctggt cgcccttggc     120
gccttctcac tatcctgtag cctgcaagcc aattcattgc cgcagatccc ggcagagcag     180
gtcaacaata ccatgtatca gatgttctac tgggatgcct atcccggact ttgggccgat     240
ctgccagcca tggcggaacc attggctgaa cgggggatca cctccatgtg gctgccgccg     300
gcggcaaaag gcatgaatgg caccaacagc gttggctacg atgtctatga cttctgggat     360
ctgggcgaat tctaccaaaa agggaccatc gccacccgtt acggcacccg ccagcagctg     420
gaccaggcgc tgagtgctct ggatcaactg ggtgttcagg cctattttga tgtggtcttt     480
aaccaccgca tgggcgccga tgcccaggaa tacattcctg gctatggact ggcctggacc     540
gagtaccagc tgcagggtcg gcaggtgcat tatacccagc aaaactgggg ctatttgtgg     600
cacgactttg actggaactg gaccgcgttt aatggctccg acaatcagtt gtatcctggt     660
aaatggtggg gcaatacctt ccacttccct tatttgatgg gcgaggatgt cgattacaac     720
cgctttgaag tgcagcagga aatgaaagcc tgggggggagt ggatcatcaa cagcgttggc     780
tttagcggct ttcggatgga tgccatcgcc catgtcgata ccgattttac ccgtgactgg     840
atcaatcacg tgcagtgggc caccagtgag gatgtgttct ttgtcgctga agcctgggtc     900
agcgatatca acggctatct ggatgccgtc aatacgccgc atttgcgcgc ttttgatttc     960
aatttgcgcg aagactttgt tgctttaagc agtggcagca aagacatgcg ctggtggggt    1020
ggtctggtca atagccagca ccgtgatcgg gcggtcacct ttgtcgataa ccacgatacc    1080
agccgggccg gcaaccctta tggcatgccg caggtgatca actacaagaa ccaggcctac    1140
gcttacattc tgttgcgtga gcatggggtg ccgaccgtgt ttgcccgcga ttacgacgaa    1200
tttggcatgg cgccaacgct ggataaactg attgaagcgc gccgttactt tgcctatggc    1260
cctggccatg agtactccgg caatactgaa gccgtctatg cctatgtgcg cgaggggcta    1320
agtactgtgc cgggcaccgg tttggtgatg ctgatgtcgg gtcgtaactg gggtggtcag    1380
cagtcgttca ccatcaacag ccaccagccg aataccacct ttacgatta caccggcaat     1440
gtcagtggca ccgtaaccac caatgcgcag ggctatggca gcttcccggt caatatgacg    1500
gaaagtaccg gttggtcagt ctgggtacca caatccattg gtggcactca gccgggaacc    1560
attaccctgc ggatgaccaa ggatgttggc tatggctttt cactgttctt caccggcagc    1620
agtgcggagc tgaccaattg gggcggcggt gttgaaggca cctggacatc cggtaatgtg    1680
tgggaagtga ccatcccgga tcca                                           1704
```

<210> 339
<211> 568
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(50)

<400> 339

```
Met Pro His Gln Tyr Asn Tyr Arg Pro Lys Ile His Arg Leu Phe Ser
 1               5                  10                  15
Gly Ala Pro Met Gln Thr Asn Lys Lys Thr Phe Leu Ala His Cys Leu
            20                  25                  30
```

```
Arg Gln Ser Leu Val Ala Leu Gly Ala Phe Ser Leu Ser Cys Ser Leu
        35                  40                  45
Gln Ala Asn Ser Leu Pro Gln Ile Pro Ala Glu Gln Val Asn Asn Thr
    50                  55                  60
Met Tyr Gln Met Phe Tyr Trp Asp Ala Tyr Pro Gly Leu Trp Ala Asp
65                  70                  75                  80
Leu Pro Ala Met Ala Glu Pro Leu Ala Glu Arg Gly Ile Thr Ser Met
                85                  90                  95
Trp Leu Pro Pro Ala Ala Lys Gly Met Asn Gly Thr Asn Ser Val Gly
            100                 105                 110
Tyr Asp Val Tyr Asp Phe Trp Asp Leu Gly Glu Phe Tyr Gln Lys Gly
        115                 120                 125
Thr Ile Ala Thr Arg Tyr Gly Thr Arg Gln Gln Leu Asp Gln Ala Leu
    130                 135                 140
Ser Ala Leu Asp Gln Leu Gly Val Gln Ala Tyr Phe Asp Val Val Phe
145                 150                 155                 160
Asn His Arg Met Gly Ala Asp Ala Gln Glu Tyr Ile Pro Gly Tyr Gly
                165                 170                 175
Leu Ala Trp Thr Glu Tyr Gln Leu Gln Gly Arg Gln Val His Tyr Thr
            180                 185                 190
Gln Gln Asn Trp Gly Tyr Leu Trp His Asp Phe Asp Trp Asn Trp Thr
        195                 200                 205
Ala Phe Asn Gly Ser Asp Asn Gln Leu Tyr Pro Gly Lys Trp Trp Gly
    210                 215                 220
Asn Thr Phe His Phe Pro Tyr Leu Met Gly Glu Asp Val Asp Tyr Asn
225                 230                 235                 240
Arg Phe Glu Val Gln Gln Glu Met Lys Ala Trp Gly Glu Trp Ile Ile
                245                 250                 255
Asn Ser Val Gly Phe Ser Gly Phe Arg Met Asp Ala Ile Ala His Val
            260                 265                 270
Asp Thr Asp Phe Thr Arg Asp Trp Ile Asn His Val Gln Trp Ala Thr
        275                 280                 285
Ser Glu Asp Val Phe Phe Val Ala Glu Ala Trp Val Ser Asp Ile Asn
    290                 295                 300
Gly Tyr Leu Asp Ala Val Asn Thr Pro His Leu Arg Ala Phe Asp Phe
305                 310                 315                 320
Asn Leu Arg Glu Asp Phe Val Ala Leu Ser Ser Gly Ser Lys Asp Met
                325                 330                 335
Arg Trp Trp Gly Gly Leu Val Asn Ser Gln His Arg Asp Arg Ala Val
            340                 345                 350
Thr Phe Val Asp Asn His Asp Thr Ser Arg Ala Gly Asn Pro Tyr Gly
        355                 360                 365
Met Pro Gln Val Ile Asn Tyr Lys Asn Gln Ala Tyr Ala Tyr Ile Leu
    370                 375                 380
Leu Arg Glu His Gly Val Pro Thr Val Phe Ala Arg Asp Tyr Asp Glu
385                 390                 395                 400
Phe Gly Met Ala Pro Thr Leu Asp Lys Leu Ile Glu Ala Arg Arg Tyr
                405                 410                 415
Phe Ala Tyr Gly Pro Gly His Glu Tyr Ser Gly Asn Thr Glu Ala Val
            420                 425                 430
Tyr Ala Tyr Val Arg Glu Gly Leu Ser Thr Val Pro Gly Thr Gly Leu
        435                 440                 445
Val Met Leu Met Ser Gly Arg Asn Trp Gly Gly Gln Gln Ser Phe Thr
    450                 455                 460
Ile Asn Ser His Gln Pro Asn Thr Thr Phe Tyr Asp Tyr Thr Gly Asn
465                 470                 475                 480
Val Ser Gly Thr Val Thr Thr Asn Ala Gln Gly Tyr Gly Ser Phe Pro
                485                 490                 495
Val Asn Met Thr Glu Ser Thr Gly Trp Ser Val Trp Val Pro Gln Ser
            500                 505                 510
Ile Gly Gly Thr Gln Pro Gly Thr Ile Thr Leu Arg Met Thr Lys Asp
        515                 520                 525
Val Gly Tyr Gly Phe Ser Leu Phe Phe Thr Gly Ser Ser Ala Glu Leu
    530                 535                 540
Thr Asn Trp Gly Gly Gly Val Glu Gly Thr Trp Thr Ser Gly Asn Val
545                 550                 555                 560
Trp Glu Val Thr Ile Pro Asp Pro
                565
```

<210> 340
<211> 1848
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 340
gtgaaacctg tcacccctac ccgcacgggc gagtcaggaa tcagctacca gttgctggtt 60
tactcgttct gcgactccga cggtgacggt accggtgatt tcaagggaat cgaaagcaag 120
ctggactatc tcaaggagat gggtgtacag gctttatggc tgtcaccaat ccatcctgca 180
acatcctatc acggctatga cgtggaggac tacagttcgg taaatccagc ctacggcacg 240
gaggcggatt tcgagtccct gctgaaagcg gcgcactcga aaggaatgaa aatctacctt 300
gatttcgtgc tgaaccacac atccaaggaa catccgtggt tccttgaggg aaagtccgat 360
ccggacagca ggtaccgaca ctattaccat ttctccacga cggccaagga cggatattcg 420
cagaccgttt cgggcaccga cccagggaaa atcaacgtga agttcaccct taaatgcaat 480
tcttccggca cacctcagac gcttaaggcc gaaagggtgg agagcgcggt caacgagggc 540
actccgaatt ccggaaaata cctctggtac ggttccgtga ccgaaagcac aatgccggaa 600
ttctattctg ccggcgagga cacatatacc cttgcaatag aaaacttcga aaccgactgg 660
ggcgtactgg tacgcaccag caagactgaa tggggcgcaa agaaattcgg cgcccaatcc 720
gagagcgccc agatgttcgt atggggcgtt ccgcttaatc tgaaatcaaa ttccgactac 780
gacatactga tgccttggat gggcaagatc tggtatcagt ccgtattcgg ctcatatatg 840
ccggacctga attacggtgc ggccgccagc tgcgaggaat cagaagcctt caaggaagtc 900
tgcgcagctg ctgacaaatg gataaaaatg ggcgttgacg gcttccgtct cgacgccgtc 960
aagcacatat atgacaacga gaactccgac gagaacccga ccttcctgaa gaaattctac 1020
gaccattgca ataccaccta taaggctgca ggtcattcag gaaatatcta tatggtcggg 1080
gagcaatgga gtgaaccgaa ttatgtcacg ccatattaca agggcctcaa cgctttcttc 1140
gagtttgcct tctgctggcg tctgaccgat gccataaacg gggccaaagg cgcaggcttc 1200
gcctccacca ttgacggtta tcaccagaaa tatactgcga cacgctccga tgccattgca 1260
gccacgaaac tctcaaacca cgacgaggac cgtatcgcct caaccctcgg ccgcaacgca 1320
agcaagctta agctggccgc ggcggtactt cttaccgccg gtggtgaacc atatatctac 1380
cagggcgaag agcttggtta ctggggcacg aaaagcaatg gggacgaata tgtccgcacc 1440
ccgatattgt ggacggctga cgcaggcagc gcagccgcgg gcgcactgaa cggcaagatc 1500
gacaggaata tgctgacgtc agacatttcc gtggaggccc aggccggcaa ccaggattct 1560
atgcttggcg tttaccgcaa attcggcgta ctccgcgact gttacaaagc tctcgcgaaa 1620
ggcagcttcc aggaaaagaa cggaatcagc aaccaggcca tctgcgcctg gtatcgcgaa 1680
tacgatggcc agaagattct tgtagtccac aatttcagcg cggcacccgt cactttcgcg 1740
ccggcaggag acaagctttc ggcaatgata ggctcaaacg gaagcgcgac ggtaaagaac 1800
gacaacttga cccttggagg ttatgcctcc gctcttttcc tccaataa 1848

<210> 341
<211> 615
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 341
Met Lys Pro Val Thr Pro Thr Arg Thr Gly Glu Ser Gly Ile Ser Tyr
1 5 10 15
Gln Leu Leu Val Tyr Ser Phe Cys Asp Ser Asp Gly Asp Gly Thr Gly
20 25 30
Asp Phe Lys Gly Ile Glu Ser Lys Leu Asp Tyr Leu Lys Glu Met Gly
35 40 45
Val Gln Ala Leu Trp Leu Ser Pro Ile His Pro Ala Thr Ser Tyr His
50 55 60
Gly Tyr Asp Val Glu Asp Tyr Ser Ser Val Asn Pro Ala Tyr Gly Thr
65 70 75 80
Glu Ala Asp Phe Glu Ser Leu Leu Lys Ala Ala His Ser Lys Gly Met
85 90 95
Lys Ile Tyr Leu Asp Phe Val Leu Asn His Thr Ser Lys Glu His Pro
100 105 110
Trp Phe Leu Glu Gly Lys Ser Asp Pro Asp Ser Arg Tyr Arg His Tyr
115 120 125
Tyr His Phe Ser Thr Thr Ala Lys Asp Gly Tyr Ser Gln Thr Val Ser

```
    130                 135                 140
Gly Thr Asp Pro Gly Lys Ile Asn Val Lys Phe Thr Leu Lys Cys Asn
145                 150                 155                 160
Ser Ser Gly Thr Pro Gln Thr Leu Lys Ala Glu Arg Val Glu Ser Ala
                165                 170                 175
Val Asn Glu Gly Thr Pro Asn Ser Gly Lys Tyr Leu Trp Tyr Gly Ser
            180                 185                 190
Val Thr Glu Ser Thr Met Pro Glu Phe Tyr Ser Ala Gly Glu Asp Thr
        195                 200                 205
Tyr Thr Leu Ala Ile Glu Asn Phe Glu Thr Asp Trp Gly Val Leu Val
    210                 215                 220
Arg Thr Ser Lys Thr Glu Trp Gly Ala Lys Lys Phe Gly Ala Gln Ser
225                 230                 235                 240
Glu Ser Ala Gln Met Phe Val Trp Gly Val Pro Leu Asn Leu Lys Ser
                245                 250                 255
Asn Ser Asp Tyr Asp Ile Leu Met Pro Trp Met Gly Lys Ile Trp Tyr
            260                 265                 270
Gln Ser Val Phe Gly Ser Tyr Met Pro Asp Leu Asn Tyr Gly Ala Ala
        275                 280                 285
Ala Ser Cys Glu Glu Ser Glu Ala Phe Lys Glu Val Cys Ala Ala Ala
    290                 295                 300
Asp Lys Trp Ile Lys Met Gly Val Asp Gly Phe Arg Leu Asp Ala Val
305                 310                 315                 320
Lys His Ile Tyr Asp Asn Glu Asn Ser Asp Glu Asn Pro Thr Phe Leu
                325                 330                 335
Lys Lys Phe Tyr Asp His Cys Asn Thr Thr Tyr Lys Ala Ala Gly His
            340                 345                 350
Ser Gly Asn Ile Tyr Met Val Gly Glu Gln Trp Ser Glu Pro Asn Tyr
        355                 360                 365
Val Thr Pro Tyr Tyr Lys Gly Leu Asn Ala Phe Phe Glu Phe Ala Phe
    370                 375                 380
Cys Trp Arg Leu Thr Asp Ala Ile Asn Gly Ala Lys Gly Ala Gly Phe
385                 390                 395                 400
Ala Ser Thr Ile Asp Gly Tyr His Gln Lys Tyr Thr Ala Thr Arg Ser
                405                 410                 415
Asp Ala Ile Ala Ala Thr Lys Leu Ser Asn His Asp Glu Asp Arg Ile
            420                 425                 430
Ala Ser Thr Leu Gly Arg Asn Ala Ser Lys Leu Lys Leu Ala Ala Ala
        435                 440                 445
Val Leu Leu Thr Ala Gly Gly Glu Pro Tyr Ile Tyr Gln Gly Glu Glu
    450                 455                 460
Leu Gly Tyr Trp Gly Thr Lys Ser Asn Gly Asp Glu Tyr Val Arg Thr
465                 470                 475                 480
Pro Ile Leu Trp Thr Ala Asp Ala Gly Ser Ala Ala Ala Gly Ala Leu
                485                 490                 495
Asn Gly Lys Ile Asp Arg Asn Met Leu Thr Ser Asp Ile Ser Val Glu
            500                 505                 510
Ala Gln Ala Gly Asn Gln Asp Ser Met Leu Gly Val Tyr Arg Lys Phe
        515                 520                 525
Gly Val Leu Arg Asp Cys Tyr Lys Ala Leu Ala Lys Gly Ser Phe Gln
    530                 535                 540
Glu Lys Asn Gly Ile Ser Asn Gln Ala Ile Cys Ala Trp Tyr Arg Glu
545                 550                 555                 560
Tyr Asp Gly Gln Lys Ile Leu Val Val His Asn Phe Ser Ala Ala Pro
                565                 570                 575
Val Thr Phe Ala Pro Ala Gly Asp Lys Leu Ser Ala Met Ile Gly Ser
            580                 585                 590
Asn Gly Ser Ala Thr Val Lys Asn Asp Asn Leu Thr Leu Gly Gly Tyr
        595                 600                 605
Ala Ser Ala Leu Phe Leu Gln
    610                 615
```

<210> 342
<211> 2061
<212> DNA
<213> Unknown

<220>

<223> Obtained from an environmental sample

<400> 342
atgaaaacac tccccaaact catgtttgtt tttttgttgc acccattaat ttctatttct 60
gcaaatgcgc aaaccacttt cgttcattta tttgaatggc agtggaatga cattgccgct 120
gaatgcgaaa acgttttggg cccaaaaggt tatgccgcag tacaagtatc gccaccgcaa 180
aaatcagttg aaggaagctt ttggtggaca cgctatcaac ctgttagtta ttcgattgaa 240
ggacgttcgg gaacacgcac cgaatttgca agtatggttt ctcgatgcaa agcagtgggt 300
gtcgatattt atgtcgatgc cgttatcaat cacatggctg ctggcaaccg caattttcct 360
gaagttccct atggcgtgaa tgactttcat acctgcaccg gcaatattga ttacaccaac 420
gcatggcaag tacaaaattg cgatctggtt ggactgaatg atttaaaaac cgaatccgat 480
tatgtacgcg gaaaaattgc cgcttatatg aatgatctgt tgagcatggg tgtggcgggt 540
ttccgcattg acgcagcgaa acatatcgcc agcgcagata tcgccaatat cgtcagtcga 600
ctaaatggca atccctatat ttttcaagaa gtcataggcg cagcgaatga acccattaca 660
ccagcgcaat acacatatat tggtgatgtt actgaattta attttagtaa tacgcttggg 720
aattatttta agggtcgcgc acctctaaaa gatttacgca atataggcgt atggagtggc 780
tggttagcca gtagtgatgc agtcactttc gtagccaacc acgacaatca acgacaaaac 840
accagtaata ttgtgactca caaagatggc atgaatgtga acaacattgc gcacgttttt 900
atgctggcat ttccttacgg gtacccaaaa gttatgtcca gttatgactg gactaatcac 960
gatcaaggcc cgccaaccac tggcgcaagc aactgtaata atggctggtt atgcgagcac 1020
cgcaatcgcg aaattgccaa tatggtgggc tttcgcaaca ccactaaaga tgccttttat 1080
ctttccaatt ggtgggacaa cggcgcgaat caaattgctt ttggtcgcgg caatttaggt 1140
tttgtggtga tcaatggcga aagcagcagc atgaacgtca ccttacaaac cggtttagcc 1200
gcgggtagtt attgcaacgt gattagtggt gattttgcta acggaacctg cacgggtgcc 1260
agcattactg ttgcagctaa tggcactgcg acatttaatg tagcggcgaa gagtgcagca 1320
gcaattcata taaacgcaaa gaatggcgat atcaccggaa acaattgtgt ttatacatcg 1380
atgaaattgc gcgggacttt taacaattgg ggcagtagcg acatgatctg caacgcgggt 1440
atttggatgg gtagcgccac cttctcgggc agcagcactg atcgctttaa gtttgatgct 1500
tacggtaatt gggtcaccaa ttggggaaac aacaataatg atttaattgc cgatgctggc 1560
ggtaatgata ttgctgtcac caccgccgct acttatgcga tcacatttaa cgatgcaaat 1620
ctacaatatt cggttacccc aaccaatagc acgagcagta tcgccaccag ttcaccggcg 1680
agctcaagca gtgtgataag ttcaagtagc tcgatgagtt caagtagcgc tacaaattcg 1740
agtgttgcat ccagctccgc tgccacaact gttagcgtga acttcacttg ccaaaatggc 1800
acaacctatt caggacaaag tgtgtatgtg gtcggcaata actccacctt aggtaattgg 1860
actccagcat ccgcaatcaa actcgagcca accagctatc caatctggac cggaacgatt 1920
caattgccga gtaacactgc ggttgaatgg aaatgcctaa aacgcgaaga aaataatgca 1980
gccgctggtg ttgtgtggga gagcggcagt aacactactt ttaataccgg tagcacgacc 2040
agtaccgttg gtgcgtttta a 2061

<210> 343
<211> 686
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(23)

<400> 343
Met Lys Thr Leu Pro Lys Leu Met Phe Val Phe Leu Leu His Pro Leu
 1   5   10   15
Ile Ser Ile Ser Ala Asn Ala Gln Thr Thr Phe Val His Leu Phe Glu
            20   25   30
Trp Gln Trp Asn Asp Ile Ala Ala Glu Cys Glu Asn Val Leu Gly Pro
        35   40   45
Lys Gly Tyr Ala Ala Val Gln Val Ser Pro Pro Gln Lys Ser Val Glu
    50   55   60
Gly Ser Phe Trp Trp Thr Arg Tyr Gln Pro Val Ser Tyr Ser Ile Glu
65   70   75   80
Gly Arg Ser Gly Thr Arg Thr Glu Phe Ala Ser Met Val Ser Arg Cys
                85   90   95
Lys Ala Val Gly Val Asp Ile Tyr Val Asp Ala Val Ile Asn His Met
            100   105   110
Ala Ala Gly Asn Arg Asn Phe Pro Glu Val Pro Tyr Gly Val Asn Asp
        115   120   125
Phe His Thr Cys Thr Gly Asn Ile Asp Tyr Thr Asn Ala Trp Gln Val

```
    130                 135                 140
Gln Asn Cys Asp Leu Val Gly Leu Asn Asp Leu Lys Thr Glu Ser Asp
145                 150                 155                 160
Tyr Val Arg Gly Lys Ile Ala Ala Tyr Met Asn Asp Leu Leu Ser Met
                165                 170                 175
Gly Val Ala Gly Phe Arg Ile Asp Ala Ala Lys His Ile Ala Ser Ala
            180                 185                 190
Asp Ile Ala Asn Ile Val Ser Arg Leu Asn Gly Asn Pro Tyr Ile Phe
        195                 200                 205
Gln Glu Val Ile Gly Ala Ala Asn Glu Pro Ile Thr Pro Ala Gln Tyr
    210                 215                 220
Thr Tyr Ile Gly Asp Val Thr Glu Phe Asn Phe Ser Asn Thr Leu Gly
225                 230                 235                 240
Asn Tyr Phe Lys Gly Arg Ala Pro Leu Lys Asp Leu Arg Asn Ile Gly
                245                 250                 255
Val Trp Ser Gly Trp Leu Ala Ser Ser Asp Ala Val Thr Phe Val Ala
            260                 265                 270
Asn His Asp Asn Gln Arg Gln Asn Thr Ser Asn Ile Val Thr His Lys
        275                 280                 285
Asp Gly Met Asn Val Asn Asn Ile Ala His Val Phe Met Leu Ala Phe
    290                 295                 300
Pro Tyr Gly Tyr Pro Lys Val Met Ser Ser Tyr Asp Trp Thr Asn His
305                 310                 315                 320
Asp Gln Gly Pro Pro Thr Thr Gly Ala Ser Asn Cys Asn Asn Gly Trp
                325                 330                 335
Leu Cys Glu His Arg Asn Arg Glu Ile Ala Asn Met Val Gly Phe Arg
            340                 345                 350
Asn Thr Thr Lys Asp Ala Phe Tyr Leu Ser Asn Trp Trp Asp Asn Gly
        355                 360                 365
Ala Asn Gln Ile Ala Phe Gly Arg Gly Asn Leu Gly Phe Val Val Ile
    370                 375                 380
Asn Gly Glu Ser Ser Ser Met Asn Val Thr Leu Gln Thr Gly Leu Ala
385                 390                 395                 400
Ala Gly Ser Tyr Cys Asn Val Ile Ser Gly Asp Phe Ala Asn Gly Thr
                405                 410                 415
Cys Thr Gly Ala Ser Ile Thr Val Ala Ala Asn Gly Thr Ala Thr Phe
            420                 425                 430
Asn Val Ala Ala Lys Ser Ala Ala Ala Ile His Ile Asn Ala Lys Asn
        435                 440                 445
Gly Asp Ile Thr Gly Asn Asn Cys Val Tyr Thr Ser Met Lys Leu Arg
    450                 455                 460
Gly Thr Phe Asn Asn Trp Gly Ser Ser Asp Met Ile Cys Asn Ala Gly
465                 470                 475                 480
Ile Trp Met Gly Ser Ala Thr Phe Ser Gly Ser Ser Thr Asp Arg Phe
                485                 490                 495
Lys Phe Asp Ala Tyr Gly Asn Trp Val Thr Asn Trp Gly Asn Asn Asn
            500                 505                 510
Asn Asp Leu Ile Ala Asp Ala Gly Gly Asn Asp Ile Ala Val Thr Thr
        515                 520                 525
Ala Ala Thr Tyr Ala Ile Thr Phe Asn Asp Ala Asn Leu Gln Tyr Ser
    530                 535                 540
Val Thr Pro Thr Asn Ser Thr Ser Ser Ile Ala Thr Ser Ser Pro Ala
545                 550                 555                 560
Ser Ser Ser Ser Val Ile Ser Ser Ser Ser Ser Met Ser Ser Ser Ser
                565                 570                 575
Ala Thr Asn Ser Ser Val Ala Ser Ser Ser Ala Ala Thr Thr Val Ser
            580                 585                 590
Val Asn Phe Thr Cys Gln Asn Gly Thr Thr Tyr Ser Gly Gln Ser Val
        595                 600                 605
Tyr Val Val Gly Asn Asn Ser Thr Leu Gly Asn Trp Thr Pro Ala Ser
    610                 615                 620
Ala Ile Lys Leu Glu Pro Thr Ser Tyr Pro Ile Trp Thr Gly Thr Ile
625                 630                 635                 640
Gln Leu Pro Ser Asn Thr Ala Val Glu Trp Lys Cys Leu Lys Arg Glu
                645                 650                 655
Glu Asn Asn Ala Ala Ala Gly Val Val Trp Glu Ser Gly Ser Asn Thr
            660                 665                 670
Thr Phe Asn Thr Gly Ser Thr Thr Ser Thr Val Gly Ala Phe
```

675 680 685

<210> 344
<211> 1980
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 344
atgaaaaaac tattaattac cttgttaatt tttcctggtt ttttattctc tttttctgct 60
gttgcacaaa ccacctttgt ccatttgttt gagtggcagt ggaatgcaat tgccagtgaa 120
tgcgaaaacc atctcggccc aaaaggttat gcggcggtac aagtatcacc gccacaaaaa 180
tctatcagcg gcaaccaatg gtggacacgc taccaaccgg tcagctattc cattgaaggc 240
cgtttaggaa cccgcgccga attcgccagc atggtttcac gctgtaacgc agtgggagtg 300
gatatttatg tagacgcagt catcaatcat atgtccgccg gcaaccgcag ttttcccgaa 360
gtacctttcg gcaccaacga ctttcatact tgcaccagcg acattgacta cagcgatgtc 420
tggtctattc aaaattgcga tctggtgggt ctgaatgatt taaaaacaga gtccgattat 480
gtacgcggaa agatcgccga ctatatgaat gatttaatca gccttggcgt aaccggtttt 540
cggattgatg cggccaagca tatgcccagc agtgatattg ccaacatcgt cagccgttta 600
aatggtgacc cttatatttt tcaggaagtg atcggtgctg gcggtgagcc aatccaaccg 660
agccagtaca cctatgtcgg tgacgtaacg gaatttaatt tttccaatac gctgggtcat 720
tattttaaag gccgagccac attgcatgag ctacgcaata tcggtgtgtg gagcggatgg 780
ctgaacagtg cagatgcgat tacgtttgtc gccaaccatg ataaccagcg gcaaaatacc 840
aataacatca tcactcacaa agacggcatg aatctaaata atattgcgca cgtttttgcg 900
ctggcctggc cttacggtta tccgaaagta atgtccagtt atgactggaa cgatcacgac 960
caaggcccgc caaccaatgc tgctaacacc tgttctaacg gttggctgtg tgaacatcgc 1020
aatcgcgaaa tcgctaatat ggttggcttc cgcaatgcca cgcaaagcgc attctatgtt 1080
tccaattggt gggacaacgg cgcgaatcaa attgcttttg gtcgcggcga tcgcggtttt 1140
gtggtgatca atggcgaaag cggcggcacc ttgaatacca ctttgcaaac tggcctggct 1200
gccggtacct attgcgatgt gatcaagggt gatttttcta acggcgtatg taccggtccg 1260
accattaccg tcggtgccga cggcaatgcg agttttgctg tcggatccaa aggtgcttcg 1320
gctattcatg ctgacgctgt agtcggtccg attaccggca atgaatgcgt ttacgacacc 1380
atgaatttac gcggcacatt taacagctgg gaatcacacg ccatgacatg tgccaatggt 1440
atctggatcg gcagcgccaa cttttccggc aacagcaatg aacgattcaa gtttgacgtc 1500
tatggtgatt ggcaaaccaa ttttggtgac gccaataacg atggtgttgc caatctcggc 1560
ggcgcggata tcgcgatagc ctctgccggt gattacacca tcacctttaa cgatgcgagc 1620
tttgcttata cctttacccc ggcaaccagc agcagctcaa gttcgaattc cagcagcaca 1680
cccggtactg tggcggtgaa tttcacttgt gaaaatggca gtacttatgt gggacaaagt 1740
gtttatgtgg tgggtaacaa tgcagccatc ggcggttggt ctccaacgaa cgcggtaaag 1800
cttaatcccg tcagttatcc cacctggacc ggcacgattc aattaccggc gaacacagct 1860
attgaatgga aatgtttgaa gcgagaagaa aataatcctg cagcgggtat tgagtgggaa 1920
gcaggtggaa ataatagtct aaatacaggg agcgatgcga atacgacggg cgggttttga 1980

<210> 345
<211> 659
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(22)

<400> 345
Met Lys Lys Leu Leu Ile Thr Leu Leu Ile Phe Pro Gly Phe Leu Phe
1 5 10 15
Ser Phe Ser Ala Val Ala Gln Thr Thr Phe Val His Leu Phe Glu Trp
20 25 30
Gln Trp Asn Ala Ile Ala Ser Glu Cys Glu Asn His Leu Gly Pro Lys
35 40 45
Gly Tyr Ala Ala Val Gln Val Ser Pro Pro Gln Lys Ser Ile Ser Gly
50 55 60
Asn Gln Trp Trp Thr Arg Tyr Gln Pro Val Ser Tyr Ser Ile Glu Gly
65 70 75 80
Arg Leu Gly Thr Arg Ala Glu Phe Ala Ser Met Val Ser Arg Cys Asn

```
                85                  90                  95
Ala Val Gly Val Asp Ile Tyr Val Asp Ala Val Ile Asn His Met Ser
            100                 105                 110
Ala Gly Asn Arg Ser Phe Pro Glu Val Pro Phe Gly Thr Asn Asp Phe
        115                 120                 125
His Thr Cys Thr Ser Asp Ile Asp Tyr Ser Asp Val Trp Ser Ile Gln
    130                 135                 140
Asn Cys Asp Leu Val Gly Leu Asn Asp Leu Lys Thr Glu Ser Asp Tyr
145                 150                 155                 160
Val Arg Gly Lys Ile Ala Asp Tyr Met Asn Asp Leu Ile Ser Leu Gly
                165                 170                 175
Val Thr Gly Phe Arg Ile Asp Ala Ala Lys His Met Pro Ser Ser Asp
            180                 185                 190
Ile Ala Asn Ile Val Ser Arg Leu Asn Gly Asp Pro Tyr Ile Phe Gln
        195                 200                 205
Glu Val Ile Gly Ala Gly Gly Glu Pro Ile Gln Pro Ser Gln Tyr Thr
    210                 215                 220
Tyr Val Gly Asp Val Thr Glu Phe Asn Phe Ser Asn Thr Leu Gly His
225                 230                 235                 240
Tyr Phe Lys Gly Arg Ala Thr Leu His Glu Leu Arg Asn Ile Gly Val
                245                 250                 255
Trp Ser Gly Trp Leu Asn Ser Ala Asp Ala Ile Thr Phe Val Ala Asn
            260                 265                 270
His Asp Asn Gln Arg Gln Asn Thr Asn Asn Ile Ile Thr His Lys Asp
        275                 280                 285
Gly Met Asn Leu Asn Asn Ile Ala His Val Phe Ala Leu Ala Trp Pro
    290                 295                 300
Tyr Gly Tyr Pro Lys Val Met Ser Ser Tyr Asp Trp Asn Asp His Asp
305                 310                 315                 320
Gln Gly Pro Pro Thr Asn Ala Ala Asn Thr Cys Ser Asn Gly Trp Leu
                325                 330                 335
Cys Glu His Arg Asn Arg Glu Ile Ala Asn Met Val Gly Phe Arg Asn
            340                 345                 350
Ala Thr Gln Ser Ala Phe Tyr Val Ser Asn Trp Trp Asp Asn Gly Ala
        355                 360                 365
Asn Gln Ile Ala Phe Gly Arg Gly Asp Arg Gly Phe Val Val Ile Asn
    370                 375                 380
Gly Glu Ser Gly Gly Thr Leu Asn Thr Thr Leu Gln Thr Gly Leu Ala
385                 390                 395                 400
Ala Gly Thr Tyr Cys Asp Val Ile Lys Gly Asp Phe Ser Asn Gly Val
                405                 410                 415
Cys Thr Gly Pro Thr Ile Thr Val Gly Ala Asp Gly Asn Ala Ser Phe
            420                 425                 430
Ala Val Gly Ser Lys Gly Ala Ser Ala Ile His Ala Asp Ala Val Val
        435                 440                 445
Gly Pro Ile Thr Gly Asn Glu Cys Val Tyr Asp Thr Met Asn Leu Arg
    450                 455                 460
Gly Thr Phe Asn Ser Trp Glu Ser His Ala Met Thr Cys Ala Asn Gly
465                 470                 475                 480
Ile Trp Ile Gly Ser Ala Asn Phe Ser Gly Asn Ser Asn Glu Arg Phe
                485                 490                 495
Lys Phe Asp Val Tyr Gly Asp Trp Gln Thr Asn Phe Gly Asp Ala Asn
            500                 505                 510
Asn Asp Gly Val Ala Asn Leu Gly Gly Ala Asp Ile Ala Ile Ala Ser
        515                 520                 525
Ala Gly Asp Tyr Thr Ile Thr Phe Asn Asp Ala Ser Phe Ala Tyr Thr
    530                 535                 540
Phe Thr Pro Ala Thr Ser Ser Ser Ser Ser Asn Ser Ser Ser Thr
545                 550                 555                 560
Pro Gly Thr Val Ala Val Asn Phe Thr Cys Glu Asn Gly Ser Thr Tyr
                565                 570                 575
Val Gly Gln Ser Val Tyr Val Val Gly Asn Asn Ala Ala Ile Gly Gly
            580                 585                 590
Trp Ser Pro Thr Asn Ala Val Lys Leu Asn Pro Val Ser Tyr Pro Thr
        595                 600                 605
Trp Thr Gly Thr Ile Gln Leu Pro Ala Asn Thr Ala Ile Glu Trp Lys
    610                 615                 620
Cys Leu Lys Arg Glu Glu Asn Asn Pro Ala Ala Gly Ile Glu Trp Glu
```

625 630 635 640
Ala Gly Gly Asn Asn Ser Leu Asn Thr Gly Ser Asp Ala Asn Thr Thr
645 650 655
Gly Gly Phe

<210> 346
<211> 1398
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 346
atgaaaaaaa cagcgggaat ctgggtgatc gcaggcatgc tgatcgcccc tctggcacac 60
gccgacgtca tactgcacgc cttcaactgg caatacagcg aagtcaccgc caaggccgat 120
ctcatcaagg ccgccggcta caagcaggtg ctcatctcgc cgcccctcaa atcatcgggc 180
aacgagtggt gggcacgcta tcaaccgcag gatctgcgcc tcatcgatac tccccttggc 240
aacaagcagg atctcgagca gctgatcgcc gccatgcagg cacgaggcat agccgtctat 300
gccgatattg tgctcaacca catggccaac gagagctgga agcgcagcga tctcaactac 360
cccggcagcg agctgctcgg ccagtatgcg gccaaccccg cctacttcga acgccagaag 420
ctgttcggcg atctcggcca gaacttgctg gccggatcgg attttcaccc cgaaggctgc 480
atcaccgact ggagcgatcc gggccacgtt cagtactggc gcctgtgtgg cggcgccggt 540
gacaaggggc tgccggatct tgatcccaac aactgggtag tgagccagca gcaggcttat 600
ctcaaggcgc tcaaggggat ggggatcaag ggctttcggg tcgatgcggt caagcacatg 660
agcgattacc agatcaatgc cgtgttcacc ccagagataa aacagggcat gcacgtcttt 720
ggcgaggtaa tcaccacggg gggcgccggc agcacggatt acgaacgctt ccttaaaccc 780
tatctcgaca gcagcggtca gggggcctat gactttccgc tgtttgcctc cctgcgcggg 840
gcgctgggtt atggcggcag catgagcctg ctggccgacc ccggtgccta cggtcaggcc 900
ctgccgggca gccgcgccgt caccttcgcc atcacccacg acattcccac caacgatggc 960
tttcgctacc agatcctgaa ccagaccgac gagaagctgg cctatgccta cctgctcggt 1020
cgcgatggtg gctcgccgct tgtctattcc gatcacggcg agacccaggc caaggatggg 1080
ctgcgctggc aggattatta ccagcgcgcg gatctcaaag gcatgatccg cttccataac 1140
gcgctgcagg gccagcccat gcaactggtc ggcagcggcg actgcttcgt gctgttcaaa 1200
cgtggcaagc aggggctggt cggtgtcaac aagtgcgagt acgagcaaga gttctggctc 1260
gataccgcca aattcgagat gaactggtat cgcaactacc gggatgtgct cgatcaaaac 1320
gccgtcatca atgtgcagag ccagtgggtg cgggtcgcca ttcccgcccg cagcgcccgg 1380
ctctggctgc aagaataa 1398

<210> 347
<211> 465
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(21)

<400> 347
Met Lys Lys Thr Ala Gly Ile Trp Val Ile Ala Gly Met Leu Ile Ala
1 5 10 15
Pro Leu Ala His Ala Asp Val Ile Leu His Ala Phe Asn Trp Gln Tyr
20 25 30
Ser Glu Val Thr Ala Lys Ala Asp Leu Ile Lys Ala Ala Gly Tyr Lys
35 40 45
Gln Val Leu Ile Ser Pro Pro Leu Lys Ser Ser Gly Asn Glu Trp Trp
50 55 60
Ala Arg Tyr Gln Pro Gln Asp Leu Arg Leu Ile Asp Thr Pro Leu Gly
65 70 75 80
Asn Lys Gln Asp Leu Glu Gln Leu Ile Ala Ala Met Gln Ala Arg Gly
85 90 95
Ile Ala Val Tyr Ala Asp Ile Val Leu Asn His Met Ala Asn Glu Ser
100 105 110
Trp Lys Arg Ser Asp Leu Asn Tyr Pro Gly Ser Glu Leu Leu Gly Gln
115 120 125

Tyr Ala Ala Asn Pro Ala Tyr Phe Glu Arg Gln Lys Leu Phe Gly Asp
130 135 140
Leu Gly Gln Asn Leu Leu Ala Gly Ser Asp Phe His Pro Glu Gly Cys
145 150 155 160
Ile Thr Asp Trp Ser Asp Pro Gly His Val Gln Tyr Trp Arg Leu Cys
165 170 175
Gly Gly Ala Gly Asp Lys Gly Leu Pro Asp Leu Asp Pro Asn Asn Trp
180 185 190
Val Val Ser Gln Gln Gln Ala Tyr Leu Lys Ala Leu Lys Gly Met Gly
195 200 205
Ile Lys Gly Phe Arg Val Asp Ala Val Lys His Met Ser Asp Tyr Gln
210 215 220
Ile Asn Ala Val Phe Thr Pro Glu Ile Lys Gln Gly Met His Val Phe
225 230 235 240
Gly Glu Val Ile Thr Thr Gly Gly Ala Gly Ser Thr Asp Tyr Glu Arg
245 250 255
Phe Leu Lys Pro Tyr Leu Asp Ser Ser Gly Gln Gly Ala Tyr Asp Phe
260 265 270
Pro Leu Phe Ala Ser Leu Arg Gly Ala Leu Gly Tyr Gly Gly Ser Met
275 280 285
Ser Leu Leu Ala Asp Pro Gly Ala Tyr Gly Gln Ala Leu Pro Gly Ser
290 295 300
Arg Ala Val Thr Phe Ala Ile Thr His Asp Ile Pro Thr Asn Asp Gly
305 310 315 320
Phe Arg Tyr Gln Ile Leu Asn Gln Thr Asp Glu Lys Leu Ala Tyr Ala
325 330 335
Tyr Leu Leu Gly Arg Asp Gly Gly Ser Pro Leu Val Tyr Ser Asp His
340 345 350
Gly Glu Thr Gln Ala Lys Asp Gly Leu Arg Trp Gln Asp Tyr Tyr Gln
355 360 365
Arg Ala Asp Leu Lys Gly Met Ile Arg Phe His Asn Ala Leu Gln Gly
370 375 380
Gln Pro Met Gln Leu Val Gly Ser Gly Asp Cys Phe Val Leu Phe Lys
385 390 395 400
Arg Gly Lys Gln Gly Leu Val Gly Val Asn Lys Cys Glu Tyr Glu Gln
405 410 415
Glu Phe Trp Leu Asp Thr Ala Lys Phe Glu Met Asn Trp Tyr Arg Asn
420 425 430
Tyr Arg Asp Val Leu Asp Gln Asn Ala Val Ile Asn Val Gln Ser Gln
435 440 445
Trp Val Arg Val Ala Ile Pro Ala Arg Ser Ala Arg Leu Trp Leu Gln
450 455 460
Glu
465

<210> 348
<211> 1488
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 348
atgtctgaac ttaatggtgt aatgatgcaa tacttccact ggtatagcgc tgctgacggc 60
agcctatggc agcaagtggc agagaaagca gaggaattag ctgcggtggg ggtcacgtcg 120
ctctggttgc cgcctgccta caagggaacc gggggcggtt atgatgtggg ctacggcgtt 180
tatgacctgt ttgatctggg tgagtttgac caaaaaggct cggtgcgcac aaagtatggc 240
actaaggatg aatacctcgc tgcgattaag caggcacaac gtttaggcat caggatctat 300
gccgatgttg tgtttaacca caagttgggt gcggatgaag aagaggaagt ggaagccacg 360
cccttcaatc ctgacaatcg caatcatacg gttggcgact atcagaaaat caaagcctgg 420
acggggttca cttttcccgg tcgaggtgac aaatactcca gcatgaagtg gcattggtgg 480
cactttgatg cgatcgacta taacgcttac aaacccgaag agaacgctat ctacttgctc 540
aagggtaagg agttcgatcg taacgttgac ctcgaaaaag gtaacttcga ttacttgatg 600
ggatgtgatt tagatatgga tcatccggaa gtgattggag agttgaaata ctggggtgaa 660
tggtacatca agacaaccaa tgtcgatggg ttccggtttg atgccgtaaa gcatgtgtct 720
gctgacttct ttcgtgaatg gttagaccac gttgaacacc atgcgggtag agatgtcttt 780
gccgtaggtg agtactggtc ctatgatgtc gaggcactac acagcttcat cgaaacaacg 840

```
aatgggaaag tcacgttgtt tgatgcccca ctccactaca actttcactg cgccagtacc     900
tcaggcaaca gttatgacat gcgccaaatc tttgacggca ccctggtgca gcaacaacct     960
gctctggcag tcaccctggt agaaaaccat gattctcaac ccctgcaatc gctggaatcc    1020
atcgttgaag ggtggttcaa gcccctggca tatgccttga ttttgctgag gcgggaaggg    1080
tatccctgta ttttttatgc tgactactac ggtgctcact acaaagatca gggcagggat    1140
ggccaggagt atgagatttg gatggacagc caccaattct tgatcgacaa aatgctggtt    1200
gctcggcaaa cctttgccta tggcgatcaa tacgattatt ttgaccacgc caatacgatc    1260
ggctggacac gcctaggcac agaggagcat cctggaggca tggcggttgt gctcagtaat    1320
ggcgatgctg gcagcaaata catggaagtt gggcagccta actgcaccta tgttgacatc    1380
acggaacata ttaatgaacc gatcgtcacc aatgccgatg gatgggctga ctttcgctgt    1440
gagcctgggt cagtatcagt ttgggttccc caagcagcgg cacaatag                 1488

<210> 349
<211> 495
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 349
Met Ser Glu Leu Asn Gly Val Met Met Gln Tyr Phe His Trp Tyr Ser
 1               5                  10                  15
Ala Ala Asp Gly Ser Leu Trp Gln Gln Val Ala Glu Lys Ala Glu Glu
            20                  25                  30
Leu Ala Ala Val Gly Val Thr Ser Leu Trp Leu Pro Pro Ala Tyr Lys
        35                  40                  45
Gly Thr Gly Gly Gly Tyr Asp Val Gly Tyr Gly Val Tyr Asp Leu Phe
    50                  55                  60
Asp Leu Gly Glu Phe Asp Gln Lys Gly Ser Val Arg Thr Lys Tyr Gly
65                  70                  75                  80
Thr Lys Asp Glu Tyr Leu Ala Ala Ile Lys Gln Ala Gln Arg Leu Gly
                85                  90                  95
Ile Arg Ile Tyr Ala Asp Val Val Phe Asn His Lys Leu Gly Ala Asp
            100                 105                 110
Glu Glu Glu Glu Val Glu Ala Thr Pro Phe Asn Pro Asp Asn Arg Asn
        115                 120                 125
His Thr Val Gly Asp Tyr Gln Lys Ile Lys Ala Trp Thr Gly Phe Thr
    130                 135                 140
Phe Pro Gly Arg Gly Asp Lys Tyr Ser Ser Met Lys Trp His Trp Trp
145                 150                 155                 160
His Phe Asp Ala Ile Asp Tyr Asn Ala Tyr Lys Pro Glu Glu Asn Ala
                165                 170                 175
Ile Tyr Leu Leu Lys Gly Lys Glu Phe Asp Arg Asn Val Asp Leu Glu
            180                 185                 190
Lys Gly Asn Phe Asp Tyr Leu Met Gly Cys Asp Leu Asp Met Asp His
        195                 200                 205
Pro Glu Val Ile Gly Glu Leu Lys Tyr Trp Gly Glu Trp Tyr Ile Lys
    210                 215                 220
Thr Thr Asn Val Asp Gly Phe Arg Phe Asp Ala Val Lys His Val Ser
225                 230                 235                 240
Ala Asp Phe Phe Arg Glu Trp Leu Asp His Val Glu His His Ala Gly
                245                 250                 255
Arg Asp Val Phe Ala Val Gly Glu Tyr Trp Ser Tyr Asp Val Glu Ala
            260                 265                 270
Leu His Ser Phe Ile Glu Thr Thr Asn Gly Lys Val Thr Leu Phe Asp
        275                 280                 285
Ala Pro Leu His Tyr Asn Phe His Cys Ala Ser Thr Ser Gly Asn Ser
    290                 295                 300
Tyr Asp Met Arg Gln Ile Phe Asp Gly Thr Leu Val Gln Gln Gln Pro
305                 310                 315                 320
Ala Leu Ala Val Thr Leu Val Glu Asn His Asp Ser Gln Pro Leu Gln
                325                 330                 335
Ser Leu Glu Ser Ile Val Glu Gly Trp Phe Lys Pro Leu Ala Tyr Ala
            340                 345                 350
Leu Ile Leu Leu Arg Arg Glu Gly Tyr Pro Cys Ile Phe Tyr Ala Asp
        355                 360                 365
Tyr Tyr Gly Ala His Tyr Lys Asp Gln Gly Arg Asp Gly Gln Glu Tyr
```

```
    370                 375                 380
Glu Ile Trp Met Asp Ser His Gln Phe Leu Ile Asp Lys Met Leu Val
385                 390                 395                 400
Ala Arg Gln Thr Phe Ala Tyr Gly Asp Gln Tyr Asp Tyr Phe Asp His
                405                 410                 415
Ala Asn Thr Ile Gly Trp Thr Arg Leu Gly Thr Glu Glu His Pro Gly
            420                 425                 430
Gly Met Ala Val Val Leu Ser Asn Gly Asp Ala Gly Ser Lys Tyr Met
        435                 440                 445
Glu Val Gly Gln Pro Asn Cys Thr Tyr Val Asp Ile Thr Glu His Ile
    450                 455                 460
Asn Glu Pro Ile Val Thr Asn Ala Asp Gly Trp Ala Asp Phe Arg Cys
465                 470                 475                 480
Glu Pro Gly Ser Val Ser Val Trp Val Pro Gln Ala Ala Ala Gln
                485                 490                 495
```

<210> 350
<211> 2001
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 350

```
atgaaacttt cacacctagc tgcagcgttg ctgctggctt cgtcgagtac tgcctctctc      60
gcggcttccc cgactacgtt tgtccattta tttgagtggc gttggagtga tgttgcgcag     120
gagtgcgaaa cattttagg gccaaatggt tttgccgcag ttcaagtgtc tccgccaaac     180
gagcacattc agggttcgca gtggtggacc cgctatcaac cagtgagtta tcaactcgag     240
agtcgtggtg gtagccgtgc cgcatttatc aacatggtgc aacgttgtaa tgctgcgggg     300
gtgtcaattt atgtcgacgc cgtgatcaac catatggcca atggcagtgg cgttggtacc     360
gccggcagta attacggcaa ccgccaatat ccaatttact caggccaaga tttccatgct     420
tcttgtgcga ttaatggctc tgattacggc aataatcgtt ggcgtgtaca gaattgcgaa     480
ctctcaggtt taccagattt aaatactggc tcgccttatg tgcaaaacac aatcgccgct     540
tacatgaatg aactgaccag cttaggggta aaaggcttcc gtttggatgc agctaagcat     600
atgtcggttg aagacattgc ggcaatccgc ggcaagctga caggcgcgcc gttaatttat     660
caagaagtga tagaccaagg tggcgaagca attagctcaa gtgagtacac aggtcagggc     720
ttagttatcg agtttaaata cagtacccag ctcggcaacg tatttaaaac tggccgttta     780
gcctctctac gcaattttgg cgaaggttgg ggcttttac caagcagtcg cgccgtagtt     840
tttgtggata accatgataa ccaacgcggc catggcggtg cgggcaatgt agtgacctac     900
aaagacggcc ggatgtatga cttagccaac gtgtttatgc tggcttaccc ctacggctat     960
acccaagtta tgtcgagtta tgactataaa ggcgataccg acgctggcgg tcctggcatt    1020
ccggtccatc aaaatggtca gttgaactgt tttggcactg actggaagtg tgagcatcgt    1080
tggagttata tcgcaggggc agtgcaattt cgcaacaaca ctgttaatga atggcgcgtc    1140
accaattggt gggacaacgg caacaaccag attgcgtttg gtcgtgctgc tgctggtttt    1200
gtcgcgatca acaaagaaag ttatgcatta aataccagct tggcaaccag tatggcgcca    1260
ggtcagtatt gtaacgtgct caaaggtaaa cgcagcagtg accaaaaatc ttgtacaggt    1320
gaggtgatta ctgtgggtgg cgacggtcgg atccaagcga atgtcgccgc ttgggatgcg    1380
tttgcgattc atcaagatag caaattgacg actggcggcg gtacgccggg cgcggattgg    1440
cagcgcactg tggtgtttat ccaagcacag actctgagtg gccaagatat gtttgtccgc    1500
ggtggtcttg atcatcaagt ggcactgaac caaaaaggtt taagttgcac cgcgacgaac    1560
ttcttgtgtg caatgcctat tcgccatcgg aatctgaaaa acgcgactac agcgccatgg    1620
aaagccaacg acaattacct cgattggtac gggcgtgaag caagtcaaag cgcaactgca    1680
gaaggtacgc cgctcgattg gaccactaat ctttggccaa atagttgggg cacattacgc    1740
acagtgccgg tggatggtta tggtcaagaa gcccttaaca cgtttggcca acattactgg    1800
atgatggatg tagatatgga ctgcagcaaa accctcaacg gttggtttga actgaaggct    1860
tttgtgaaaa acggtcaagg ttgggaaggc gatatcagcc aagcaaatcc accttacgcg    1920
accaaaaacc atttagcgca atgcggcaaa ctaaacaagt ttgaatttgg cagctcggcg    1980
gcattcatca ccaacttgta a                                              2001
```

<210> 351
<211> 666
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(21)

<400> 351

```
Met Lys Leu Ser His Leu Ala Ala Ala Leu Leu Leu Ala Ser Ser Ser
 1               5                  10                  15
Thr Ala Ser Leu Ala Ala Ser Pro Thr Thr Phe Val His Leu Phe Glu
            20                  25                  30
Trp Arg Trp Ser Asp Val Ala Gln Glu Cys Glu Thr Phe Leu Gly Pro
        35                  40                  45
Asn Gly Phe Ala Ala Val Gln Val Ser Pro Pro Asn Glu His Ile Gln
    50                  55                  60
Gly Ser Gln Trp Trp Thr Arg Tyr Gln Pro Val Ser Tyr Gln Leu Glu
65                  70                  75                  80
Ser Arg Gly Gly Ser Arg Ala Ala Phe Ile Asn Met Val Gln Arg Cys
                85                  90                  95
Asn Ala Ala Gly Val Ser Ile Tyr Val Asp Ala Val Ile Asn His Met
            100                 105                 110
Ala Asn Gly Ser Gly Val Gly Thr Ala Gly Ser Asn Tyr Gly Asn Arg
        115                 120                 125
Gln Tyr Pro Ile Tyr Ser Gly Gln Asp Phe His Ala Ser Cys Ala Ile
    130                 135                 140
Asn Gly Ser Asp Tyr Gly Asn Asn Arg Trp Arg Val Gln Asn Cys Glu
145                 150                 155                 160
Leu Ser Gly Leu Pro Asp Leu Asn Thr Gly Ser Pro Tyr Val Gln Asn
                165                 170                 175
Thr Ile Ala Ala Tyr Met Asn Glu Leu Thr Ser Leu Gly Val Lys Gly
            180                 185                 190
Phe Arg Leu Asp Ala Ala Lys His Met Ser Val Glu Asp Ile Ala Ala
        195                 200                 205
Ile Arg Gly Lys Leu Thr Gly Ala Pro Leu Ile Tyr Gln Glu Val Ile
    210                 215                 220
Asp Gln Gly Gly Glu Ala Ile Ser Ser Ser Glu Tyr Thr Gly Gln Gly
225                 230                 235                 240
Leu Val Ile Glu Phe Lys Tyr Ser Thr Gln Leu Gly Asn Val Phe Lys
                245                 250                 255
Thr Gly Arg Leu Ala Ser Leu Arg Asn Phe Gly Glu Gly Trp Gly Phe
            260                 265                 270
Leu Pro Ser Ser Arg Ala Val Val Phe Val Asp Asn His Asp Asn Gln
        275                 280                 285
Arg Gly His Gly Gly Ala Gly Asn Val Val Thr Tyr Lys Asp Gly Arg
    290                 295                 300
Met Tyr Asp Leu Ala Asn Val Phe Met Leu Ala Tyr Pro Tyr Gly Tyr
305                 310                 315                 320
Thr Gln Val Met Ser Ser Tyr Asp Tyr Lys Gly Asp Thr Asp Ala Gly
                325                 330                 335
Gly Pro Gly Ile Pro Val His Gln Asn Gly Gln Leu Asn Cys Phe Gly
            340                 345                 350
Thr Asp Trp Lys Cys Glu His Arg Trp Ser Tyr Ile Ala Gly Ala Val
        355                 360                 365
Gln Phe Arg Asn Asn Thr Val Asn Glu Trp Arg Val Thr Asn Trp Trp
    370                 375                 380
Asp Asn Gly Asn Asn Gln Ile Ala Phe Gly Arg Ala Ala Ala Gly Phe
385                 390                 395                 400
Val Ala Ile Asn Lys Glu Ser Tyr Ala Leu Asn Thr Ser Leu Ala Thr
                405                 410                 415
Ser Met Ala Pro Gly Gln Tyr Cys Asn Val Leu Lys Gly Lys Arg Ser
            420                 425                 430
Ser Asp Gln Lys Ser Cys Thr Gly Glu Val Ile Thr Val Gly Gly Asp
        435                 440                 445
Gly Arg Ile Gln Ala Asn Val Ala Ala Trp Asp Ala Phe Ala Ile His
    450                 455                 460
Gln Asp Ser Lys Leu Thr Thr Gly Gly Gly Thr Pro Gly Ala Asp Trp
465                 470                 475                 480
Gln Arg Thr Val Val Phe Ile Gln Ala Gln Thr Leu Ser Gly Gln Asp
                485                 490                 495
Met Phe Val Arg Gly Gly Leu Asp His Gln Val Ala Leu Asn Gln Lys
            500                 505                 510
```

```
Gly Leu Ser Cys Thr Ala Thr Asn Phe Leu Cys Ala Met Pro Ile Arg
        515                 520                 525
His Arg Asn Leu Lys Asn Ala Thr Thr Ala Pro Trp Lys Ala Asn Asp
    530                 535                 540
Asn Tyr Leu Asp Trp Tyr Gly Arg Glu Ala Ser Gln Ser Ala Thr Ala
545                 550                 555                 560
Glu Gly Thr Pro Leu Asp Trp Thr Thr Asn Leu Trp Pro Asn Ser Trp
                565                 570                 575
Gly Thr Leu Arg Thr Val Pro Val Asp Gly Tyr Gly Gln Glu Ala Leu
            580                 585                 590
Asn Thr Phe Gly Gln His Tyr Trp Met Met Asp Val Asp Met Asp Cys
        595                 600                 605
Ser Lys Thr Leu Asn Gly Trp Phe Glu Leu Lys Ala Phe Val Lys Asn
    610                 615                 620
Gly Gln Gly Trp Glu Gly Asp Ile Ser Gln Ala Asn Pro Pro Tyr Ala
625                 630                 635                 640
Thr Lys Asn His Leu Ala Gln Cys Gly Lys Leu Asn Lys Phe Glu Phe
                645                 650                 655
Gly Ser Ser Ala Ala Phe Ile Thr Asn Leu
            660                 665
```

<210> 352
<211> 1263
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 352

```
atgctcaaga agaaggtgct tggcggagtt ctgtccctgt cgtgggccct gggtctggtg      60
gcgtcgacgg ccgcgtccgc gggtccgctg gatggcaaca gcagcgacgt gatgctgcag     120
ggcttccact ggcggtccac cgagacgtac ccctggtggg gcgtcatcca gggcaaggcc     180
tcggacatca aggccagcgg cttcaccatg gtgtggttcc caccctccgg ggactcggcg     240
gccaatgagg gctacctgcc gcggcagctc tacgtgcaga gcagtaagta cggcacggac     300
gcgcagctca agagcgccat cggctcgctg cacacctacg gggtgaaggc catcgcggac     360
atcgtggtca accaccgcgt gggcaccacc aactgggcgg acttcaccaa ccccacctgg     420
ggctcgtggt cggtggtgaa gggcgacgag tggacgagcg ccaccggcaa ctacgacacc     480
ggtgatggct tcagcgccgc gcgcgacctg gaccacacga acggcaccgt gcagagcgac     540
ctgaagggct ggatgaactg gctcaagtcc agcatcggct acgacggctg gcgctatgac     600
tacgtgaagg gctacggcgg ctcgtacgtg ggcggctaca acagcgccac ggtgccctac     660
ttctccgtgg gcgagctgtg gacggacctc aacctcaacg acgtcaaccc gcaccgccag     720
ctcatcatga actggatcaa cgccaccggg ggcaactcgg gcgcgttcga cttcaccacc     780
aagggcatcc tccagcaggc ggtgcagtac aacgagttct ggcggctcaa ggccagcgat     840
ggaaagcccc agggcgccat cggctggtgg ccggccaggt cggtgacgtt catcgacaac     900
cacgacaccg gcccgagcca cccgagcggc ggccagaacc actggccctt ccccagcgac     960
aaggtgatgc agggctacgc gtacatcctc acgcacccgg ggattccctg tgtctactgg    1020
gtgcacttct atgactgggg ccacgcagcg gccatcaaga gcctcatcgc ggcgcgcaag    1080
gagaagggcg tcacctccac gtccgcggtg aacatcgtgg ccgcggacac gagcaagtac    1140
gccgccatca tcaccggcaa cacgggcagc ctggcgatga agattggtgt cggcgcctgg    1200
tcgccgggca cgggctggac gctggtgaac tccggcacca actgggccgt ctggaagcag    1260
tag                                                                  1263
```

<210> 353
<211> 420
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(27)

<400> 353

```
Met Leu Lys Lys Lys Val Leu Gly Gly Val Leu Ser Leu Ser Trp Ala
 1               5                  10                  15
Leu Gly Leu Val Ala Ser Thr Ala Ala Ser Ala Gly Pro Leu Asp Gly
```

```
            20                  25                  30
Asn Ser Ser Asp Val Met Leu Gln Gly Phe His Trp Arg Ser Thr Glu
        35                  40                  45
Thr Tyr Pro Trp Trp Gly Val Ile Gln Gly Lys Ala Ser Asp Ile Lys
    50                  55                  60
Ala Ser Gly Phe Thr Met Val Trp Phe Pro Pro Ser Gly Asp Ser Ala
65                  70                  75                  80
Ala Asn Glu Gly Tyr Leu Pro Arg Gln Leu Tyr Val Gln Ser Ser Lys
                85                  90                  95
Tyr Gly Thr Asp Ala Gln Leu Lys Ser Ala Ile Gly Ser Leu His Thr
            100                 105                 110
Tyr Gly Val Lys Ala Ile Ala Asp Ile Val Val Asn His Arg Val Gly
        115                 120                 125
Thr Thr Asn Trp Ala Asp Phe Thr Asn Pro Thr Trp Gly Ser Trp Ser
    130                 135                 140
Val Val Lys Gly Asp Glu Trp Thr Ser Ala Thr Gly Asn Tyr Asp Thr
145                 150                 155                 160
Gly Asp Gly Phe Ser Ala Ala Arg Asp Leu Asp His Thr Asn Gly Thr
                165                 170                 175
Val Gln Ser Asp Leu Lys Gly Trp Met Asn Trp Leu Lys Ser Ser Ile
            180                 185                 190
Gly Tyr Asp Gly Trp Arg Tyr Asp Tyr Val Lys Gly Tyr Gly Gly Ser
        195                 200                 205
Tyr Val Gly Gly Tyr Asn Ser Ala Thr Val Pro Tyr Phe Ser Val Gly
    210                 215                 220
Glu Leu Trp Thr Asp Leu Asn Leu Asn Asp Val Asn Pro His Arg Gln
225                 230                 235                 240
Leu Ile Met Asn Trp Ile Asn Ala Thr Gly Gly Asn Ser Gly Ala Phe
                245                 250                 255
Asp Phe Thr Thr Lys Gly Ile Leu Gln Gln Ala Val Gln Tyr Asn Glu
            260                 265                 270
Phe Trp Arg Leu Lys Ala Ser Asp Gly Lys Pro Gln Gly Ala Ile Gly
        275                 280                 285
Trp Trp Pro Ala Arg Ser Val Thr Phe Ile Asp Asn His Asp Thr Gly
    290                 295                 300
Pro Ser His Pro Ser Gly Gly Gln Asn His Trp Pro Phe Pro Ser Asp
305                 310                 315                 320
Lys Val Met Gln Gly Tyr Ala Tyr Ile Leu Thr His Pro Gly Ile Pro
                325                 330                 335
Cys Val Tyr Trp Val His Phe Tyr Asp Trp Gly His Ala Ala Ala Ile
            340                 345                 350
Lys Ser Leu Ile Ala Ala Arg Lys Glu Lys Gly Val Thr Ser Thr Ser
        355                 360                 365
Ala Val Asn Ile Val Ala Ala Asp Thr Ser Lys Tyr Ala Ala Ile Ile
    370                 375                 380
Thr Gly Asn Thr Gly Ser Leu Ala Met Lys Ile Gly Val Gly Ala Trp
385                 390                 395                 400
Ser Pro Gly Thr Gly Trp Thr Leu Val Asn Ser Gly Thr Asn Trp Ala
                405                 410                 415
Val Trp Lys Gln
            420
```

<210> 354
<211> 2577
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 354

```
atgcacaaca caatgtttcg aaccgccatg ctcgccgcgg tgctgggctc attcagccac      60
acagccgccg ccgaaggggt catggtccac ctgttccaat ggaagttcaa cgatatcgcc     120
aacgagtgcg agacagtgct cggtcccaag ggattcggtg gtgtccagat cactccgccc     180
gcagagcaca agcaggggag ccaggtctgg tggaccgtct accagcccgt cagcttcaag     240
aacttcaaca gcttcggcgg cagcgaggca gagctcagga gcatgatcac ccgctgcaat     300
gcggccgggg tcaaggtcta tgccgacgcc gtcttcaacc agctggcctc gggcagcggg     360
acagccaccg gcggcggcag ctacaactcg gggcaatatc agtatcccca gttcggctac     420
```

```
aacgacttcc accacagcgg tgacatcacc aactacggcg acagcaacaa cgtctggaac     480
ggcgccctct acggcatgcc ggatctcaat accggctctc cctacgtgca ggatcagata     540
gccacctaca tgaagaccct gctgggttgg ggcgtggcgg gctttcgcat cgatgcggcc     600
aagcacatgg cccccggcga ggtgaaggcc atactcgaca aggcgggcag ccccaaggcc     660
tacctggaag tgataggcgc aggtggagag tcggcagaca tccagccggg tcgctacacc     720
tatatcgaca ctgtcaccga cttcaaatac ggcacagatc tggccgccaa cttcaacggc     780
cagatcaaga acctcaagac cctgggcgag agctggggcc tgctcccctc cagcaaggcc     840
ttcgtcttcg tggtcaacca cgaccgcgag cgcggccacg gtggcggcgg catgctcacc     900
ttcatgagcg gtgcccgcta cgatctggcc aacaccttca tgatggcctg gccctatggc     960
tggaagcagg tgatgtcggg ctatcgcttc gagaacatgg gcacctatga aaccgacaag    1020
ggggcgccgg gcagcacacc ctgcacagac ggccagtgga actgcgagca gcgtcgcccg    1080
accatcatga acatggccct gttccacaac agaaccgagg gtcaagctgt cagtcactgg    1140
tgggacaacg gcaacaacca gatagccttc ggccggggag acaaaggctt cgtcgccatc    1200
aacaacgaga gcggcagcct ggtggcctcg gtgcagaccg gcctgccagc gggcgagtac    1260
tgcaacctcc tcggcggcaa tgactactgc agcgggggtt atgtcaccgt cgacggcagc    1320
ggcaaggcca gcctgaacgt gccgggcatg aaggccgccg ccatcatcgc cggctgcacc    1380
aaggccaacc cctgcggcgg caccgccctg ccgggcaaca agttcagcag catgaacctg    1440
cgtggcaccc acaacgcctg gggcaacacc cccatgacag tcgatgccaa ccgtgtctgg    1500
tccgccaccc tgaccctgac cggcagcggt gacgccaccg gtgcccagcg tttcaagttc    1560
gacgtcttcg gcaactggac cgagagctac ggcgacaacg aaggggacgg catcgccgac    1620
aagggcagca gcaaggatat cctggtcagc gggacgggca gctaccgcat caccctgaac    1680
gaaggggatc tgcgctatac ggtcaccccc ctcaccagca accaggcacc ggtcgccgcc    1740
ctctcgccca aggtcctcag cgtcaagacg ggagagagcg tggtattcga tggctccgcc    1800
tccaccgacg atgaagccgt ggcgagctac agctggtcta ccggcggcaa cgcccccact    1860
gagaccgtcc agttcgatac cccgggcacc cataccgtca ccctgacagt caccgatgcc    1920
gaggggctga cctccagtgc cagcgcaact gtgactgtca ccgacagcaa cggcgcctac    1980
aacagcgtgc tgccgaccct gcacttccgt ggcaccccca atggttgggg gacgctcgcc    2040
atgacgctgg ttgccgacaa ccagtgggag gcattggcca ccttcaacgg ccaggccaac    2100
cagcgcttca agttcgacgt caagggagac tggacccaga actacggtga caccaacaag    2160
gatggggtgg ccgaacaggg gggcgccgac atagtgaccc cggtcagcgg ccagtaccgg    2220
gtgcgcttca acgaccagac cctgcaatac agcctgaccc cggtcagcgt cggctatgcc    2280
aagaacatcg ccagcctgaa catccgtggc accaccaatg gctggggcac cactcccatg    2340
acgctggtgg gggatcacca gtggcaggcc ggtgtcacct tcacgggggc cggtgatgcc    2400
agcggcggcc agcgcttcaa gttcgacgtc aagggtgact ggacccagaa ctacggcgac    2460
accaacaagg acggcgtggc cgagctggcc ggagccgaca tcaccaccgc agtggtcggt    2520
gcctatgtgg tgcgcttcaa cgaccagacc ctggcctaca gcctgagcgc ccagtaa       2577

<210> 355
<211> 858
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(24)

<400> 355
Met His Asn Thr Met Phe Arg Thr Ala Met Leu Ala Ala Val Leu Gly
 1               5                  10                  15
Ser Phe Ser His Thr Ala Ala Ala Glu Gly Val Met Val His Leu Phe
            20                  25                  30
Gln Trp Lys Phe Asn Asp Ile Ala Asn Glu Cys Glu Thr Val Leu Gly
        35                  40                  45
Pro Lys Gly Phe Gly Gly Val Gln Ile Thr Pro Pro Ala Glu His Lys
    50                  55                  60
Gln Gly Ser Gln Val Trp Trp Thr Val Tyr Gln Pro Val Ser Phe Lys
65                  70                  75                  80
Asn Phe Asn Ser Phe Gly Gly Ser Glu Ala Glu Leu Arg Ser Met Ile
                85                  90                  95
Thr Arg Cys Asn Ala Ala Gly Val Lys Val Tyr Ala Asp Ala Val Phe
            100                 105                 110
Asn Gln Leu Ala Ser Gly Ser Gly Thr Ala Thr Gly Gly Gly Ser Tyr
        115                 120                 125
Asn Ser Gly Gln Tyr Gln Tyr Pro Gln Phe Gly Tyr Asn Asp Phe His
    130                 135                 140
His Ser Gly Asp Ile Thr Asn Tyr Gly Asp Ser Asn Asn Val Trp Asn
```

```
145                 150                 155                 160
Gly Ala Leu Tyr Gly Met Pro Asp Leu Asn Thr Gly Ser Pro Tyr Val
                165                 170                 175
Gln Asp Gln Ile Ala Thr Tyr Met Lys Thr Leu Leu Gly Trp Gly Val
            180                 185                 190
Ala Gly Phe Arg Ile Asp Ala Ala Lys His Met Ala Pro Gly Glu Val
        195                 200                 205
Lys Ala Ile Leu Asp Lys Ala Gly Ser Pro Lys Ala Tyr Leu Glu Val
    210                 215                 220
Ile Gly Ala Gly Gly Glu Ser Ala Asp Ile Gln Pro Gly Arg Tyr Thr
225                 230                 235                 240
Tyr Ile Asp Thr Val Thr Asp Phe Lys Tyr Gly Thr Asp Leu Ala Ala
                245                 250                 255
Asn Phe Asn Gly Gln Ile Lys Asn Leu Lys Thr Leu Gly Glu Ser Trp
            260                 265                 270
Gly Leu Leu Pro Ser Ser Lys Ala Phe Val Phe Val Val Asn His Asp
        275                 280                 285
Arg Glu Arg Gly His Gly Gly Gly Gly Met Leu Thr Phe Met Ser Gly
    290                 295                 300
Ala Arg Tyr Asp Leu Ala Asn Thr Phe Met Met Ala Trp Pro Tyr Gly
305                 310                 315                 320
Trp Lys Gln Val Met Ser Gly Tyr Arg Phe Glu Asn Met Gly Thr Tyr
                325                 330                 335
Glu Thr Asp Lys Gly Ala Pro Gly Ser Thr Pro Cys Thr Asp Gly Gln
            340                 345                 350
Trp Asn Cys Glu Gln Arg Arg Pro Thr Ile Met Asn Met Ala Leu Phe
        355                 360                 365
His Asn Arg Thr Glu Gly Gln Ala Val Ser His Trp Trp Asp Asn Gly
    370                 375                 380
Asn Asn Gln Ile Ala Phe Gly Arg Gly Asp Lys Gly Phe Val Ala Ile
385                 390                 395                 400
Asn Asn Glu Ser Gly Ser Leu Val Ala Ser Val Gln Thr Gly Leu Pro
                405                 410                 415
Ala Gly Glu Tyr Cys Asn Leu Leu Gly Gly Asn Asp Tyr Cys Ser Gly
            420                 425                 430
Gly Tyr Val Thr Val Asp Gly Ser Gly Lys Ala Ser Leu Asn Val Pro
        435                 440                 445
Gly Met Lys Ala Ala Ala Ile Ile Ala Gly Cys Thr Lys Ala Asn Pro
    450                 455                 460
Cys Gly Gly Thr Ala Leu Pro Gly Asn Lys Phe Ser Ser Met Asn Leu
465                 470                 475                 480
Arg Gly Thr His Asn Ala Trp Gly Asn Thr Pro Met Thr Val Asp Ala
                485                 490                 495
Asn Arg Val Trp Ser Ala Thr Leu Thr Leu Thr Gly Ser Gly Asp Ala
            500                 505                 510
Thr Gly Ala Gln Arg Phe Lys Phe Asp Val Phe Gly Asn Trp Thr Glu
        515                 520                 525
Ser Tyr Gly Asp Asn Glu Gly Asp Gly Ile Ala Asp Lys Gly Ser Ser
    530                 535                 540
Lys Asp Ile Leu Val Ser Gly Thr Gly Ser Tyr Arg Ile Thr Leu Asn
545                 550                 555                 560
Glu Gly Asp Leu Arg Tyr Thr Val Thr Pro Leu Thr Ser Asn Gln Ala
                565                 570                 575
Pro Val Ala Ala Leu Ser Pro Lys Val Leu Ser Val Lys Thr Gly Glu
            580                 585                 590
Ser Val Val Phe Asp Gly Ser Ala Ser Thr Asp Asp Glu Ala Val Ala
        595                 600                 605
Ser Tyr Ser Trp Ser Thr Gly Gly Asn Ala Pro Thr Glu Thr Val Gln
    610                 615                 620
Phe Asp Thr Pro Gly Thr His Thr Val Thr Leu Thr Val Thr Asp Ala
625                 630                 635                 640
Glu Gly Leu Thr Ser Ser Ala Ser Ala Thr Val Thr Val Thr Asp Ser
                645                 650                 655
Asn Gly Ala Tyr Asn Ser Val Leu Pro Thr Leu His Phe Arg Gly Thr
            660                 665                 670
Pro Asn Gly Trp Gly Thr Leu Ala Met Thr Leu Val Ala Asp Asn Gln
        675                 680                 685
Trp Glu Ala Leu Ala Thr Phe Asn Gly Gln Ala Asn Gln Arg Phe Lys
```

```
    690                 695                 700
Phe Asp Val Lys Gly Asp Trp Thr Gln Asn Tyr Gly Asp Thr Asn Lys
705                 710                 715                 720
Asp Gly Val Ala Glu Gln Gly Gly Ala Asp Ile Val Thr Pro Val Ser
                725                 730                 735
Gly Gln Tyr Arg Val Arg Phe Asn Asp Gln Thr Leu Gln Tyr Ser Leu
            740                 745                 750
Thr Pro Val Ser Val Gly Tyr Ala Lys Asn Ile Ala Ser Leu Asn Ile
        755                 760                 765
Arg Gly Thr Thr Asn Gly Trp Gly Thr Thr Pro Met Thr Leu Val Gly
    770                 775                 780
Asp His Gln Trp Gln Ala Gly Val Thr Phe Thr Gly Ala Gly Asp Ala
785                 790                 795                 800
Ser Gly Gly Gln Arg Phe Lys Phe Asp Val Lys Gly Asp Trp Thr Gln
                805                 810                 815
Asn Tyr Gly Asp Thr Asn Lys Asp Gly Val Ala Glu Leu Ala Gly Ala
            820                 825                 830
Asp Ile Thr Thr Ala Val Val Gly Ala Tyr Val Val Arg Phe Asn Asp
        835                 840                 845
Gln Thr Leu Ala Tyr Ser Leu Ser Ala Gln
    850                 855
```

<210> 356
<211> 4875
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 356

```
gtgatgctgg ttttacaatg gtatggttcc ccccccgtat ctgaatcagt ggataaccat      60
gggtatcttc ccggaaggct gtatgacctt aacagctcaa aatacggcaa ccagcagcag     120
ctgattgagg ctatccaggc gcttaatgca gcggatgtta aagccatggc tgatattgtt     180
atcaatcacc gctgcggaca gtatcagatt aacggtaact ggcacggata tgaacagccg     240
gactggcagg aatgggctgt tgttgctaat gatggcggca cgggtgctca ggataccggt     300
gagggatacg gtcctgcacc ggacatagac cacaccaacc ctacagtcca ggcggacatt     360
atggcctgga tgaactggct gcggactaca ataggatttg atggctggcg ctatgattat     420
gtcagaggct tttccggtta ctatgtgggt atgtacaacg atgccacaca accgtatttt     480
tcagtcgggg aattctggcc agatattaca gggagttatt atgctactgt ccctgacttg     540
aactaccacc ggcagcagct tgctgactgg gtggatgcaa caggtaatgc atcgcacgtt     600
attgatttta caaccaagtg gcagctccag cttgcttttg agcggacgga attctggcgg     660
atgggacagg ttcccggttt gattggctgg tggccggagc gaagtgtgac gtttattgac     720
aatcatgata caggttccag cccgggcgga gggcagtcgc attgggagtt tcctgcggtg     780
cacgtggcgg caggctatgc gtatattctt acacatcccg gtgtaccaaa ggtgtattgg     840
ccgcattatt ttgattgggg ctcagggctg agaaatacta tcagagacct catggcaatc     900
agactaaata acgggataca ttcgggcagc accattgatg tgcaggttgc tgaagcctcc     960
agatatgctg ccattattga cgggaaggtt gccgtcaaaa ttggcccggg tgactggtca    1020
ccgggcgggg actggcacct tgcagcgtcc gggaatggct ttgctgtttg gaccaaggat    1080
gaaggcggct ctgatcgtca atatgatacc atgttcatca ggggcacgtt taataactgg    1140
ggctcaactg caatgactcc gcttggggat tcagagtgga ttgcggtggc tgaattttcc    1200
ggtgcttcca atgaccgctt caagtttgat gcctacggtg actggtcttt gaactggggt    1260
gacaataacg gagacggcgt actcaatcaa tttggagagg atattccaat tactgaggga    1320
gccggcagat ataccataac gctcaatgaa gaaactatgg cgtatactgt gacgtttgat    1380
ggcagtattg atattccagc ggcaccttca ggtttaagcg ctcaggcaat ttcttctgca    1440
gctatccgtg tatcctgggc accagtgtca ggtgctgctt cctatacagt gtacagggct    1500
gcatctgctg atgggacata cactgctgta ggtacaacag gatcaacgaa ctttgatgtt    1560
tccggtttat tcagcgacag gacatattat ttccgcgttg ctgcttccaa tgcagccggc    1620
aactctgaac aatctgcata tgtttcggct accacccca aagccacaac gattgttcgc    1680
gcgttctata cggctgccgg cggcaaccgc ttatatatac ggggtgatat ggctccgcta    1740
aactggaata ccggtgaaga aatgacccgc gtgtctgacg gtttatggga atggcagact    1800
gatgcgcttt ccacggcacg aaatgttaat tttaaaattc ttatcaatga cagtatctgg    1860
tcagaaggtg cgatgttcag ggtgattggc gggcagacac tggaaattca tcccagtttt    1920
aatcagcctc cggcaatacc tcaggggctg gtaggtcagg ctctctcacc tttttcagta    1980
cggctcaact ggcatgccgt ttccggcgcg gacaggtata ttttgtatca atccactgtt    2040
tcatggggaa attatatagt agctgcgact actgaggaaa ggcaggtaga cattgagggg    2100
ctggatgctg aggctacgta ttatttcagg gttgcggctg aaaatgacta cggcacctcc    2160
ccccgttcca cgtttattgc tgttactacg cccagcgctg tgcctgctgt accggcggga    2220
```

```
ctcgaagcag aacctgtttc aggcacgagc atcgctcttt cctggggttc tgttaccggg    2280
gctgagacgt atacagtgta ccgtgcagca gctgctgacg gtacatatgc tgccatcggt    2340
actacaagtt cggcgaattt tgatgcaaca ggacttcatt cagaccggac atacttttтc    2400
cgcgtggatg catctaatga atttggaaga tctgcgcaat cctcccctgt ttctgcaagt    2460
acaccggcgg tagtaacgac tgtacgtgcc cactttgcgt cagctcaggg agagacctta    2520
tcgctcaggg gtgatatggc accgcttaac tgggatactg gtgcagcaat gtcgtatctt    2580
ggtgatggtg tctggcaatg gcagactgat gagctttcaa ctgcgagaaa tattaactac    2640
aaagttctgc tcaatgatac aatatggtca gaaggtgcta tgtaccggat tattggcgga    2700
cagattgtgg atatttatcc gtcgtttaat cagccgcccg ctattccctc cggcctgcag    2760
gcagagacgc tttcttcatc agaaatagca atacactggg atgcggtttc agatgcggac    2820
tcgtatacag tatatcgttc cacggttaac tggggtggat atgaggcaat cggtaccact    2880
gcggatactg aatttcaggc agtggggctt gccagctcta ctacgtattt cttccgggtc    2940
agcgctgcta atgcatatgg ggcatccccc cgttctgtgt ttgtctcggc aactacggaa    3000
ggggaggttc ccggtattcc ctctagtgtg gctgcggaag ccttgtctgc atctgcaata    3060
aaggtgagct gggataccgc caccggtgct gacgggtatg aggtgtatag atccatctcc    3120
tcagatggtc cgtttacgat tgtggcatcg accagcgaaa ctacgcacat cgaaaacggc    3180
cttgcttcag aaactgtata ctatttcagg gtcaacgcag ttaatgaatg ggggaactcc    3240
ggattctctt caacagccag cgcatcaaca cttgggacag ctccttcagc tccctcaggt    3300
ctgagtttga cggcattgtc cggaacctca gtcagggctc agtgggatgc agtttccggt    3360
gcggaatatt atactgttta ccgttccagc gcggcagacg ggtcgtatac agctgccgga    3420
acggcgagct ctccatctat ggatatttcc agcctgcata gtgatagaac atatttttтc    3480
cgggttagcg cctccaacca gtggggtgag tcaagtcaat catcacatgc ttcagtgagc    3540
acaccggcgg tcactacgac tgtaagggcg cattattcgc cttcccaggg tgagggattg    3600
tatttgagag gtgatatggc acctctgaac tggaatacag gtgtagcaat gacctatacc    3660
ggcgaaggta tctgggaatg gcaaacggat gccctgtcta cagcgcgtaa tattaactat    3720
aaggtgctcc ttgatgatgc aacctggtcc gaggggggcaa tgtaccgggt aataggcggc    3780
cagactgttg atatttatcc tgggtttaat acacctcctg ctattccgca gaatgtttct    3840
gccgatacac agtcttcctc tacggttctg gtaagctggg atgccgtgga tgatacagag    3900
tactacaggg tttttatgtc caccgtttcc tggggagaat attccgaaat aggtgttacc    3960
ggtgatacca gcctgcaggc tgagggtctt gacagcctca caacatatta tttcagggta    4020
agtgcggtaa atgagtatgg cacatctgcg cgttctacat ttgtgtcggc aaccactttt    4080
ggtgtagccc cttctgtacc ggctaatatc agtgcctctc cgctaaccgc ttcttccata    4140
cgtgtaagct gggatgctgt ttccgcagtt gaggaatacc gcgtgtacat ggcggaatct    4200
tccggtggcg aatatcagac agccggattg acatctaatt catcttttga cgtttccggc    4260
ctcaacccgg caaccacata ttattttaaa gtccgttctt tgaatgagtg gggctattca    4320
aaccaatcca gtgctgtttc cgctgctacc tatgcccttc ctgccaatat cccaacaggc    4380
cttaccgctg aggctgtttc ttcatcggca atcaatctgg tatgggatgc tgttgaatat    4440
gcagaaacat atacagtata ccgttcaact tcagaagatg gaacgtatac atctattggc    4500
actactgaat ccccgtcttt tcaggcggat aatttaagtg tttccacgcg gtactatttt    4560
aaagtacgcg cggttaatgc ccacggcact actgagcagt ctgttgctgt gacagcaaga    4620
acatttggaa cgcctccttc tgttcctcaa aatgttgaag cagttgtttt atcttctaca    4680
agaatcattg tgagttggaa ttctgttgca gctgctgacg attatattgt ttttatcgcg    4740
gctgattccg ccggccccta cgaggaggca ggcattattt cctctacggt atatgaggct    4800
gaaaatttgt cagccaaaac aacgtattat ttcagagttg ctgcggttaa tgagtggggt    4860
acttccagcg cttct                                                     4875

<210> 357
<211> 1625
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 357
Met Met Leu Val Leu Gln Trp Tyr Gly Ser Pro Pro Val Ser Glu Ser
 1               5                  10                  15
Val Asp Asn His Gly Tyr Leu Pro Gly Arg Leu Tyr Asp Leu Asn Ser
            20                  25                  30
Ser Lys Tyr Gly Asn Gln Gln Gln Leu Ile Glu Ala Ile Gln Ala Leu
        35                  40                  45
Asn Ala Ala Asp Val Lys Ala Met Ala Asp Ile Val Ile Asn His Arg
    50                  55                  60
Cys Gly Gln Tyr Gln Ile Asn Gly Asn Trp His Gly Tyr Glu Gln Pro
65                  70                  75                  80
Asp Trp Gln Glu Trp Ala Val Val Ala Asn Asp Gly Gly Thr Gly Ala
                85                  90                  95
Gln Asp Thr Gly Glu Gly Tyr Gly Pro Ala Pro Asp Ile Asp His Thr
```

```
                100                 105                 110
Asn Pro Thr Val Gln Ala Asp Ile Met Ala Trp Met Asn Trp Leu Arg
        115                 120                 125
Thr Thr Ile Gly Phe Asp Gly Trp Arg Tyr Asp Tyr Val Arg Gly Phe
    130                 135                 140
Ser Gly Tyr Tyr Val Gly Met Tyr Asn Asp Ala Thr Gln Pro Tyr Phe
145                 150                 155                 160
Ser Val Gly Glu Phe Trp Pro Asp Ile Thr Gly Ser Tyr Tyr Ala Thr
                165                 170                 175
Val Pro Asp Leu Asn Tyr His Arg Gln Gln Leu Ala Asp Trp Val Asp
            180                 185                 190
Ala Thr Gly Asn Ala Ser His Val Ile Asp Phe Thr Thr Lys Trp Gln
        195                 200                 205
Leu Gln Leu Ala Phe Glu Arg Thr Glu Phe Trp Arg Met Gly Gln Val
    210                 215                 220
Pro Gly Leu Ile Gly Trp Trp Pro Glu Arg Ser Val Thr Phe Ile Asp
225                 230                 235                 240
Asn His Asp Thr Gly Ser Ser Pro Gly Gly Gly Gln Ser His Trp Glu
                245                 250                 255
Phe Pro Ala Val His Val Ala Ala Gly Tyr Ala Tyr Ile Leu Thr His
            260                 265                 270
Pro Gly Val Pro Lys Val Tyr Trp Pro His Tyr Phe Asp Trp Gly Ser
        275                 280                 285
Gly Leu Arg Asn Thr Ile Arg Asp Leu Met Ala Ile Arg Leu Asn Asn
    290                 295                 300
Gly Ile His Ser Gly Ser Thr Ile Asp Val Gln Val Ala Glu Ala Ser
305                 310                 315                 320
Arg Tyr Ala Ala Ile Ile Asp Gly Lys Val Ala Val Lys Ile Gly Pro
                325                 330                 335
Gly Asp Trp Ser Pro Gly Gly Asp Trp His Leu Ala Ala Ser Gly Asn
            340                 345                 350
Gly Phe Ala Val Trp Thr Lys Asp Glu Gly Gly Ser Asp Arg Gln Tyr
        355                 360                 365
Asp Thr Met Phe Ile Arg Gly Thr Phe Asn Asn Trp Gly Ser Thr Ala
    370                 375                 380
Met Thr Pro Leu Gly Asp Ser Glu Trp Ile Ala Val Ala Glu Phe Ser
385                 390                 395                 400
Gly Ala Ser Asn Asp Arg Phe Lys Phe Asp Ala Tyr Gly Asp Trp Ser
                405                 410                 415
Leu Asn Trp Gly Asp Asn Asn Gly Asp Gly Val Leu Asn Gln Phe Gly
            420                 425                 430
Glu Asp Ile Pro Ile Thr Glu Gly Ala Gly Arg Tyr Thr Ile Thr Leu
        435                 440                 445
Asn Glu Glu Thr Met Ala Tyr Thr Val Thr Phe Asp Gly Ser Ile Asp
    450                 455                 460
Ile Pro Ala Ala Pro Ser Gly Leu Ser Ala Gln Ala Ile Ser Ser Ala
465                 470                 475                 480
Ala Ile Arg Val Ser Trp Ala Pro Val Ser Gly Ala Ala Ser Tyr Thr
                485                 490                 495
Val Tyr Arg Ala Ala Ser Ala Asp Gly Thr Tyr Thr Ala Val Gly Thr
            500                 505                 510
Thr Gly Ser Thr Asn Phe Asp Val Ser Gly Leu Phe Ser Asp Arg Thr
        515                 520                 525
Tyr Tyr Phe Arg Val Ala Ala Ser Asn Ala Ala Gly Asn Ser Glu Gln
    530                 535                 540
Ser Ala Tyr Val Ser Ala Thr Thr Pro Lys Ala Thr Thr Ile Val Arg
545                 550                 555                 560
Ala Phe Tyr Thr Ala Ala Gly Gly Asn Arg Leu Tyr Ile Arg Gly Asp
                565                 570                 575
Met Ala Pro Leu Asn Trp Asn Thr Gly Glu Glu Met Thr Arg Val Ser
            580                 585                 590
Asp Gly Leu Trp Glu Trp Gln Thr Asp Ala Leu Ser Thr Ala Arg Asn
        595                 600                 605
Val Asn Phe Lys Ile Leu Ile Asn Asp Ser Ile Trp Ser Glu Gly Ala
    610                 615                 620
Met Phe Arg Val Ile Gly Gly Gln Thr Leu Glu Ile His Pro Ser Phe
625                 630                 635                 640
Asn Gln Pro Pro Ala Ile Pro Gln Gly Leu Val Gly Gln Ala Leu Ser
```

645 650 655
Pro Phe Ser Val Arg Leu Asn Trp His Ala Val Ser Gly Ala Asp Arg
660 665 670
Tyr Ile Leu Tyr Gln Ser Thr Val Ser Trp Gly Asn Tyr Ile Val Ala
675 680 685
Ala Thr Thr Glu Glu Arg Gln Val Asp Ile Glu Gly Leu Asp Ala Glu
690 695 700
Ala Thr Tyr Tyr Phe Arg Val Ala Ala Glu Asn Asp Tyr Gly Thr Ser
705 710 715 720
Pro Arg Ser Thr Phe Ile Ala Val Thr Thr Pro Ser Ala Val Pro Ala
725 730 735
Val Pro Ala Gly Leu Glu Ala Glu Pro Val Ser Gly Thr Ser Ile Ala
740 745 750
Leu Ser Trp Gly Ser Val Thr Gly Ala Glu Thr Tyr Thr Val Tyr Arg
755 760 765
Ala Ala Ala Ala Asp Gly Thr Tyr Ala Ala Ile Gly Thr Thr Ser Ser
770 775 780
Ala Asn Phe Asp Ala Thr Gly Leu His Ser Asp Arg Thr Tyr Phe Phe
785 790 795 800
Arg Val Asp Ala Ser Asn Glu Phe Gly Arg Ser Ala Gln Ser Ser Pro
805 810 815
Val Ser Ala Ser Thr Pro Ala Val Val Thr Thr Val Arg Ala His Phe
820 825 830
Ala Ser Ala Gln Gly Glu Thr Leu Ser Leu Arg Gly Asp Met Ala Pro
835 840 845
Leu Asn Trp Asp Thr Gly Ala Ala Met Ser Tyr Leu Gly Asp Gly Val
850 855 860
Trp Gln Trp Gln Thr Asp Glu Leu Ser Thr Ala Arg Asn Ile Asn Tyr
865 870 875 880
Lys Val Leu Leu Asn Asp Thr Ile Trp Ser Glu Gly Ala Met Tyr Arg
885 890 895
Ile Ile Gly Gly Gln Ile Val Asp Ile Tyr Pro Ser Phe Asn Gln Pro
900 905 910
Pro Ala Ile Pro Ser Gly Leu Gln Ala Glu Thr Leu Ser Ser Ser Glu
915 920 925
Ile Ala Ile His Trp Asp Ala Val Ser Asp Ala Asp Ser Tyr Thr Val
930 935 940
Tyr Arg Ser Thr Val Asn Trp Gly Gly Tyr Glu Ala Ile Gly Thr Thr
945 950 955 960
Ala Asp Thr Glu Phe Gln Ala Val Gly Leu Ala Ser Ser Thr Thr Tyr
965 970 975
Phe Phe Arg Val Ser Ala Ala Asn Ala Tyr Gly Ala Ser Pro Arg Ser
980 985 990
Val Phe Val Ser Ala Thr Thr Glu Gly Glu Val Pro Gly Ile Pro Ser
995 1000 1005
Ser Val Ala Ala Glu Ala Leu Ser Ala Ser Ala Ile Lys Val Ser Trp
1010 1015 1020
Asp Thr Ala Thr Gly Ala Asp Gly Tyr Glu Val Tyr Arg Ser Ile Ser
1025 1030 1035 1040
Ser Asp Gly Pro Phe Thr Ile Val Ala Ser Thr Ser Glu Thr Thr His
1045 1050 1055
Ile Glu Asn Gly Leu Ala Ser Glu Thr Val Tyr Tyr Phe Arg Val Asn
1060 1065 1070
Ala Val Asn Glu Trp Gly Asn Ser Gly Phe Ser Ser Thr Ala Ser Ala
1075 1080 1085
Ser Thr Leu Gly Thr Ala Pro Ser Ala Pro Ser Gly Leu Ser Leu Thr
1090 1095 1100
Ala Leu Ser Gly Thr Ser Val Arg Ala Gln Trp Asp Ala Val Ser Gly
1105 1110 1115 1120
Ala Glu Tyr Tyr Thr Val Tyr Arg Ser Ser Ala Ala Asp Gly Ser Tyr
1125 1130 1135
Thr Ala Ala Gly Thr Ala Ser Ser Pro Ser Met Asp Ile Ser Ser Leu
1140 1145 1150
His Ser Asp Arg Thr Tyr Phe Phe Arg Val Ser Ala Ser Asn Gln Trp
1155 1160 1165
Gly Glu Ser Ser Gln Ser Ser His Ala Ser Val Ser Thr Pro Ala Val
1170 1175 1180
Thr Thr Thr Val Arg Ala His Tyr Ser Pro Ser Gln Gly Glu Gly Leu

```
1185                1190                1195                1200
Tyr Leu Arg Gly Asp Met Ala Pro Leu Asn Trp Asn Thr Gly Val Ala
            1205                1210                1215
Met Thr Tyr Thr Gly Glu Gly Ile Trp Glu Trp Gln Thr Asp Ala Leu
        1220                1225                1230
Ser Thr Ala Arg Asn Ile Asn Tyr Lys Val Leu Leu Asp Asp Ala Thr
    1235                1240                1245
Trp Ser Glu Gly Ala Met Tyr Arg Val Ile Gly Gly Gln Thr Val Asp
1250                1255                1260
Ile Tyr Pro Gly Phe Asn Thr Pro Pro Ala Ile Pro Gln Asn Val Ser
1265                1270                1275                1280
Ala Asp Thr Gln Ser Ser Ser Thr Val Leu Val Ser Trp Asp Ala Val
            1285                1290                1295
Asp Asp Thr Glu Tyr Tyr Arg Val Phe Met Ser Thr Val Ser Trp Gly
        1300                1305                1310
Glu Tyr Ser Glu Ile Gly Val Thr Gly Asp Thr Ser Leu Gln Ala Glu
    1315                1320                1325
Gly Leu Asp Ser Leu Thr Thr Tyr Tyr Phe Arg Val Ser Ala Val Asn
1330                1335                1340
Glu Tyr Gly Thr Ser Ala Arg Ser Thr Phe Val Ser Ala Thr Thr Phe
1345                1350                1355                1360
Gly Val Ala Pro Ser Val Pro Ala Asn Ile Ser Ala Ser Pro Leu Thr
            1365                1370                1375
Ala Ser Ser Ile Arg Val Ser Trp Asp Ala Val Ser Ala Val Glu Glu
        1380                1385                1390
Tyr Arg Val Tyr Met Ala Glu Ser Ser Gly Gly Glu Tyr Gln Thr Ala
    1395                1400                1405
Gly Leu Thr Ser Asn Ser Ser Phe Asp Val Ser Gly Leu Asn Pro Ala
1410                1415                1420
Thr Thr Tyr Tyr Phe Lys Val Arg Ser Leu Asn Glu Trp Gly Tyr Ser
1425                1430                1435                1440
Asn Gln Ser Ser Ala Val Ser Ala Ala Thr Tyr Ala Leu Pro Ala Asn
            1445                1450                1455
Ile Pro Thr Gly Leu Thr Ala Glu Ala Val Ser Ser Ser Ala Ile Asn
        1460                1465                1470
Leu Val Trp Asp Ala Val Glu Tyr Ala Glu Thr Tyr Thr Val Tyr Arg
    1475                1480                1485
Ser Thr Ser Glu Asp Gly Thr Tyr Thr Ser Ile Gly Thr Thr Glu Ser
1490                1495                1500
Pro Ser Phe Gln Ala Asp Asn Leu Ser Val Ser Thr Arg Tyr Tyr Phe
1505                1510                1515                1520
Lys Val Arg Ala Val Asn Ala His Gly Thr Thr Glu Gln Ser Val Ala
            1525                1530                1535
Val Thr Ala Arg Thr Phe Gly Thr Pro Pro Ser Val Pro Gln Asn Val
        1540                1545                1550
Glu Ala Val Val Leu Ser Ser Thr Arg Ile Ile Val Ser Trp Asn Ser
    1555                1560                1565
Val Ala Ala Ala Asp Asp Tyr Ile Val Phe Ile Ala Ala Asp Ser Ala
1570                1575                1580
Gly Pro Tyr Glu Glu Ala Gly Ile Ile Ser Ser Thr Val Tyr Glu Ala
1585                1590                1595                1600
Glu Asn Leu Ser Ala Lys Thr Thr Tyr Tyr Phe Arg Val Ala Ala Val
            1605                1610                1615
Asn Glu Trp Gly Thr Ser Ser Ala Ser
        1620                1625

<210> 358
<211> 1422
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 358
atgcaagtaa tcagtactgt tgtaaaacgc cggctgagct ggctgttgtt tgtgctaacg      60
ccgctgctga gcagtaccgc cagcgccgat gccattttac atgcctttga ctggcattat     120
gacgacgtgg ctgccaaagc taccgagatc aaagatctgg gctataaagc ggtgctggtc     180
```

```
gcaccgccgt taaagtctaa tgcggctaac tgtgcctggt ggcagcgcta tcagccgcag     240
gatatccgtg taatagatca ttgcaaaggc aacaagcaag cctttgtaaa tatgatcaac     300
gcgttaaacg atgctaaccc ggcgcgcaag gtcgacgtgt atgctgatat agtgttaaac     360
catatggcta atgagcgcag cggtgccact gattttcccg gtagtgcagc agtaaatagc     420
tatggcagca acagcagcta ctggaataac cagcagctgt ttggcaattt aactcaggga     480
attttcagcg ccggcgactt taacccggct aactgcatca gcaattataa cgatgtctgg     540
caggtacaaa attaccgctt atgtggtggc gctggcgata ccggcttgcc ggatttaaac     600
ccgaacagct gggtagtgca gcagcaacgc agttacttga cggcactgaa aaacctcggc     660
gtaaagggct ttcgtgtcga tgcggccaag catatgacca tttggcatat caacgaaatt     720
ttcaccagca gcattaaaag cggtagctac ctgtttggtg aaattattac cagcggtggc     780
gctggcaaca acgagtacga cagtttttta gcgccttacc tgaactatac cgatcataaa     840
gcctatgatt ttccgctgtt taacagtatt cgtaatgcct tcagttttgg tggcagttta     900
agccagctgg taaacccggc ggcttatggc caggcgctgc aaaacagccg ggccgttacc     960
tttaccatta cgcatgatat tccgaccaat gatggctttc gctatctgat catggacccg    1020
accgatgagt atctggccta tgcctatgtt atgggccgcg atggcggtaa gccgctgatc    1080
ttcagtgaca gcaccggcac ggacaataac cgctgggtca atgcttataa agcggatcat    1140
attagcaaga tgttgaattt tcataaccgc atgcagggcc agggcatgga aatgctggcg    1200
tggaatgact gcgctattct gtttcgccgt ggccaggaag gcatagtagg catcaataaa    1260
tgcagtggca gccagagttt tagtgttaat acccatggcc gcttttactg gtaccgcaat    1320
taccgcgatg tgttaaccgg tggcaatctg gtgtacatca acggcggtaa ctatcacttc    1380
agcatacccg cgcgtggcgc gcggatgtgg tatgccgact aa                       1422

<210> 359
<211> 473
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(29)

<400> 359
Met Gln Val Ile Ser Thr Val Val Lys Arg Arg Leu Ser Trp Leu Leu
 1               5                  10                  15
Phe Val Leu Thr Pro Leu Leu Ser Ser Thr Ala Ser Ala Asp Ala Ile
            20                  25                  30
Leu His Ala Phe Asp Trp His Tyr Asp Asp Val Ala Ala Lys Ala Thr
        35                  40                  45
Glu Ile Lys Asp Leu Gly Tyr Lys Ala Val Leu Val Ala Pro Pro Leu
    50                  55                  60
Lys Ser Asn Ala Ala Asn Cys Ala Trp Trp Gln Arg Tyr Gln Pro Gln
65                  70                  75                  80
Asp Ile Arg Val Ile Asp His Cys Lys Gly Asn Lys Gln Ala Phe Val
                85                  90                  95
Asn Met Ile Asn Ala Leu Asn Asp Ala Asn Pro Ala Arg Lys Val Asp
            100                 105                 110
Val Tyr Ala Asp Ile Val Leu Asn His Met Ala Asn Glu Arg Ser Gly
        115                 120                 125
Ala Thr Asp Phe Pro Gly Ser Ala Ala Val Asn Ser Tyr Gly Ser Asn
    130                 135                 140
Ser Ser Tyr Trp Asn Asn Gln Gln Leu Phe Gly Asn Leu Thr Gln Gly
145                 150                 155                 160
Ile Phe Ser Ala Gly Asp Phe Asn Pro Ala Asn Cys Ile Ser Asn Tyr
                165                 170                 175
Asn Asp Val Trp Gln Val Gln Asn Tyr Arg Leu Cys Gly Gly Ala Gly
            180                 185                 190
Asp Thr Gly Leu Pro Asp Leu Asn Pro Asn Ser Trp Val Val Gln Gln
        195                 200                 205
Gln Arg Ser Tyr Leu Thr Ala Leu Lys Asn Leu Gly Val Lys Gly Phe
    210                 215                 220
Arg Val Asp Ala Ala Lys His Met Thr Ile Trp His Ile Asn Glu Ile
225                 230                 235                 240
Phe Thr Ser Ser Ile Lys Ser Gly Ser Tyr Leu Phe Gly Glu Ile Ile
                245                 250                 255
Thr Ser Gly Gly Ala Gly Asn Asn Glu Tyr Asp Ser Phe Leu Ala Pro
            260                 265                 270
```

```
Tyr Leu Asn Tyr Thr Asp His Lys Ala Tyr Asp Phe Pro Leu Phe Asn
        275                 280                 285
Ser Ile Arg Asn Ala Phe Ser Phe Gly Gly Ser Leu Ser Gln Leu Val
    290                 295                 300
Asn Pro Ala Ala Tyr Gly Gln Ala Leu Gln Asn Ser Arg Ala Val Thr
305                 310                 315                 320
Phe Thr Ile Thr His Asp Ile Pro Thr Asn Asp Gly Phe Arg Tyr Leu
                325                 330                 335
Ile Met Asp Pro Thr Asp Glu Tyr Leu Ala Tyr Ala Tyr Val Met Gly
            340                 345                 350
Arg Asp Gly Gly Lys Pro Leu Ile Phe Ser Asp Ser Thr Gly Thr Asp
        355                 360                 365
Asn Asn Arg Trp Val Asn Ala Tyr Lys Ala Asp His Ile Ser Lys Met
    370                 375                 380
Leu Asn Phe His Asn Arg Met Gln Gly Gln Gly Met Glu Met Leu Ala
385                 390                 395                 400
Trp Asn Asp Cys Ala Ile Leu Phe Arg Arg Gly Gln Glu Gly Ile Val
                405                 410                 415
Gly Ile Asn Lys Cys Ser Gly Ser Gln Ser Phe Ser Val Asn Thr His
            420                 425                 430
Gly Arg Phe Tyr Trp Tyr Arg Asn Tyr Arg Asp Val Leu Thr Gly Gly
        435                 440                 445
Asn Leu Val Tyr Ile Asn Gly Gly Asn Tyr His Phe Ser Ile Pro Ala
    450                 455                 460
Arg Gly Ala Arg Met Trp Tyr Ala Asp
465                 470
```

<210> 360
<211> 1938
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 360

```
ttgagtgtaa taaattgtga agaagtgaaa cgagatgagt ttcatacaga aaagtactat      60
gaaagttata acatctttgg cgcacatgtt gtgacggaag atgagattca aggggtacga     120
tttacagtct gggctcctca tgcgaaagca atgagtgttg ttggagattt taatgagtgg     180
gattatgagc aacataagat gctacaagtg acagaagaag ggatttggtc cttatttata     240
ccgcatattg aagaaggaga aatatataaa tatgcgattg aaacgttggc tggtgacgtc     300
attttaaagg cagatccgta tgctatatat gcagaagtaa gaccgaatac ggcatctgta     360
gtttttgata taaaaggata tgaatggaat gataaaaact ggaatcgtaa gaaaaagaaa     420
aaaccgattt ataaagaagc gatgacagtt tatgaattac attttggttc ttggaaaaag     480
aaagaagatg gaacactata ttcttacagg gaaatggcag aagaactcat tccgtatgtg     540
gtggaacatc aatttacaca tattgaaatt atgccgcttg ttgagcatcc atatgatcgt     600
tcttggggat atcaaggaac gggatattat gcagcgacga gtagatttgg tacaccgcat     660
gatttaatgt attttgtcga cgaatgtcat aaatatggaa tcggtgtcat tttagattgg     720
gtaccggggc atttttgtaa agatgctcac ggtttatatt tgtttgatgg cacaccgacc     780
tatgaatata aagataaaga tgtacaagaa aatctagtat ggggaactgt caattttgat     840
ttaggaaaga gagaggtacg taatttctta atttcaaatg cgttattttg gatgagatat     900
ttccatattg atggtttcag agtagatgca gttgcgaaca tgttgtactg gaataaagaa     960
ggacaagagc aaagtaatga gtacgctgtg tcatttttaa gggagttaaa tgaagcagtg    1020
tttgcagagg atgaagattt tcttatgacg gcagaggatt caacagcttg gccacttgta    1080
acagcaccaa cgtatgaagg tgggcttgga ttcaattaca aatggaatat gggctggatg    1140
aatgacgtgc tgaaatatat ggagtgtgcg ccagagtaca gaaaatatat tcatgagaaa    1200
atgacgttct ctttactata tgcttactct gaaaacttca tattacccct ttctcatgat    1260
gaagtcgttc atgggaaaaa gtcgttatta aataaaatgc caggtgatta ctgggataag    1320
tttgctcagc ttcgtttatt atatggatat ttctttactc acccaggaaa gaaattactt    1380
ttcatgggag gagaattcgg gcagtttgat gagtggaaag accttgaaga tttagattgg    1440
aatttacatg attttgaaat gcatcgttat atgcatgatt actttaaaga gctcatagca    1500
ttgtataagc gctcaaaacc actttggcag cttgatcatt cgcctgaagg ttttcagtgg    1560
attgatgcta ataataacga gcaaagtatt ttctcgttta tccgccaagg ggataaacaa    1620
gaagatgcgt tagttatcgt atgtaatttt acgaaagcta catatgaaaa ctataaagta    1680
ggtgtaccag atttcgagta ttataacgag attttaaaca gtgatgcaca gcaatatggc    1740
ggttcgggac aagtgaataa gaaacgtctt aagacgattc tagaaccgta ccataatcaa    1800
gcagcacatg tagagattac aattccacca tttggcgtat ccatattacg accagtgaag    1860
acgagaaagg ggagcaaaaa acaagatggc tcaaaaacaa aagtgcgtag caatgttact    1920
```

agcaggggga aaaggtag 1938

<210> 361
<211> 645
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 361

```
Leu Ser Val Ile Asn Cys Glu Glu Val Lys Arg Asp Glu Phe His Thr
 1               5                  10                  15
Glu Lys Tyr Tyr Glu Ser Tyr Asn Ile Phe Gly Ala His Val Val Thr
            20                  25                  30
Glu Asp Glu Ile Gln Gly Val Arg Phe Thr Val Trp Ala Pro His Ala
        35                  40                  45
Lys Ala Met Ser Val Val Gly Asp Phe Asn Glu Trp Asp Tyr Glu Gln
    50                  55                  60
His Lys Met Leu Gln Val Thr Glu Glu Gly Ile Trp Ser Leu Phe Ile
65                  70                  75                  80
Pro His Ile Glu Glu Gly Glu Ile Tyr Lys Tyr Ala Ile Glu Thr Leu
                85                  90                  95
Ala Gly Asp Val Ile Leu Lys Ala Asp Pro Tyr Ala Ile Tyr Ala Glu
            100                 105                 110
Val Arg Pro Asn Thr Ala Ser Val Val Phe Asp Ile Lys Gly Tyr Glu
        115                 120                 125
Trp Asn Asp Lys Asn Trp Asn Arg Lys Lys Lys Lys Lys Pro Ile Tyr
    130                 135                 140
Lys Glu Ala Met Thr Val Tyr Glu Leu His Phe Gly Ser Trp Lys Lys
145                 150                 155                 160
Lys Glu Asp Gly Thr Leu Tyr Ser Tyr Arg Glu Met Ala Glu Glu Leu
                165                 170                 175
Ile Pro Tyr Val Val Glu His Gln Phe Thr His Ile Glu Ile Met Pro
            180                 185                 190
Leu Val Glu His Pro Tyr Asp Arg Ser Trp Gly Tyr Gln Gly Thr Gly
        195                 200                 205
Tyr Tyr Ala Ala Thr Ser Arg Phe Gly Thr Pro His Asp Leu Met Tyr
    210                 215                 220
Phe Val Asp Glu Cys His Lys Tyr Gly Ile Gly Val Ile Leu Asp Trp
225                 230                 235                 240
Val Pro Gly His Phe Cys Lys Asp Ala His Gly Leu Tyr Leu Phe Asp
                245                 250                 255
Gly Thr Pro Thr Tyr Glu Tyr Lys Asp Lys Asp Val Gln Glu Asn Leu
            260                 265                 270
Val Trp Gly Thr Val Asn Phe Asp Leu Gly Lys Arg Glu Val Arg Asn
        275                 280                 285
Phe Leu Ile Ser Asn Ala Leu Phe Trp Met Arg Tyr Phe His Ile Asp
    290                 295                 300
Gly Phe Arg Val Asp Ala Val Ala Asn Met Leu Tyr Trp Asn Lys Glu
305                 310                 315                 320
Gly Gln Glu Gln Ser Asn Glu Tyr Ala Val Ser Phe Leu Arg Glu Leu
                325                 330                 335
Asn Glu Ala Val Phe Ala Glu Asp Glu Asp Phe Leu Met Thr Ala Glu
            340                 345                 350
Asp Ser Thr Ala Trp Pro Leu Val Thr Ala Pro Thr Tyr Glu Gly Gly
        355                 360                 365
Leu Gly Phe Asn Tyr Lys Trp Asn Met Gly Trp Met Asn Asp Val Leu
    370                 375                 380
Lys Tyr Met Glu Cys Ala Pro Glu Tyr Arg Lys Tyr Ile His Glu Lys
385                 390                 395                 400
Met Thr Phe Ser Leu Leu Tyr Ala Tyr Ser Glu Asn Phe Ile Leu Pro
                405                 410                 415
Leu Ser His Asp Glu Val Val His Gly Lys Lys Ser Leu Leu Asn Lys
            420                 425                 430
Met Pro Gly Asp Tyr Trp Asp Lys Phe Ala Gln Leu Arg Leu Leu Tyr
        435                 440                 445
Gly Tyr Phe Phe Thr His Pro Gly Lys Lys Leu Leu Phe Met Gly Gly
```

```
    450                 455                 460
Glu Phe Gly Gln Phe Asp Glu Trp Lys Asp Leu Glu Asp Leu Asp Trp
465                 470                 475                 480
Asn Leu His Asp Phe Glu Met His Arg Tyr Met His Asp Tyr Phe Lys
                485                 490                 495
Glu Leu Ile Ala Leu Tyr Lys Arg Ser Lys Pro Leu Trp Gln Leu Asp
            500                 505                 510
His Ser Pro Glu Gly Phe Gln Trp Ile Asp Ala Asn Asn Asn Glu Gln
        515                 520                 525
Ser Ile Phe Ser Phe Ile Arg Gln Gly Asp Lys Gln Glu Asp Ala Leu
    530                 535                 540
Val Ile Val Cys Asn Phe Thr Lys Ala Thr Tyr Glu Asn Tyr Lys Val
545                 550                 555                 560
Gly Val Pro Asp Phe Glu Tyr Tyr Asn Glu Ile Leu Asn Ser Asp Ala
                565                 570                 575
Gln Gln Tyr Gly Gly Ser Gly Gln Val Asn Lys Lys Arg Leu Lys Thr
            580                 585                 590
Ile Leu Glu Pro Tyr His Asn Gln Ala Ala His Val Glu Ile Thr Ile
        595                 600                 605
Pro Pro Phe Gly Val Ser Ile Leu Arg Pro Val Lys Thr Arg Lys Gly
    610                 615                 620
Ser Lys Lys Gln Asp Gly Ser Lys Thr Lys Val Arg Ser Asn Val Thr
625                 630                 635                 640
Ser Arg Gly Lys Arg
                645
```

<210> 362
<211> 2094
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 362

```
atgaaaaaaa ttactttctt attggcatta ttgcttttgg taagcgcaac aatatttgcc      60
aacgacccca aagacgacat aatgctgcaa tcatttggtt gggacgaata cgaccagccc     120
aaaataaagg ctgctggtaa cttctacaaa tttgttgagt cgaaagctag cgagtggaaa     180
gctgctggtt tcgatatgat atggatgcca ccccctagtc gctccacagg tggtgtgggt     240
tatttgccca ccgagttaca caacttcaac tcaacatggg gcagcgaggc cgacttacgt     300
gctgctacgg ctgcactcaa agcacatggc atgtttccta ttgccgatgt ggttgccaat     360
catcgcaatg gcaccacggc atggaccgac tttacgaacc ccgcatggag ctgtgcaacc     420
atcacatcta ccgacgaggc taatgtaaat tgggttgcat cgcgtggccc aaagccttgt     480
ggcgcacccg atacaggcga cgattttaat ggcgggcgcg acctcgacca taccaatatc     540
gaaacacgca acggcataaa aacttttctc acaaaactaa aagaacaagg cttcgaaggc     600
tggcgttggg acatgaccaa aggttttagt gccagttatg taggcgagta taacactgca     660
agccaaccct atttctcggt aggcgaatat tgggacggaa actcgaacac cctcaaaaac     720
tgggtagatg gcacaggcaa aaagtcggca acattcgatt tttcgcaata ttattcgatg     780
gagaatgcat tcaaaaataa tgcatggagt gcattaggtt cgggaagcag catggccggt     840
ttggcgggtg tttttggcta ctccgactat gccgttactt ttgtcgacaa ccacgatact     900
tttgtgcatg gcagcgcgcc actcggcgat aatattatga aagcctatgc ctacatactt     960
acgcaccccg gcataccgag cgtatttatt gcgcattatt acggtggaac atatcgtaaa    1020
gatggtgtta ctcgcactta cacaagccac aaagataaaa tcgaccccat aatggcagtg    1080
cgcaaagcca atcgtattga cgcttggagc acattacaag ttgcaacaaa ttcgggcatg    1140
tatgccgctt atatcaaacg ccgccacgac gatgccgctg cttcggtagc cgtaaaaatt    1200
ggcccaggaa gttggtcgcc cgaaggtagt ggctggattt tagctgcatc gggaacagac    1260
tacgccgtgt ggtcgaaagc gcaaatagtt gtagcaccta ccatgctcat tggtagtgga    1320
acttacgaat taggacaaac actcagcatt acagctccgt cgggctatag catacgctac    1380
accaccgatg gtagcgagcc aacagcggcg tcgacacttt acactgcgcc cattacattg    1440
ccacaaggca caaccgttat caaagctgct tcgtttgtaa atggtgttat gtcggccgta    1500
gtaaccaaca cttacactgc acaagcaaaa gcaaccagca taaaagtgcg tttcaaagct    1560
cccgcttcgt ggacagcaag caaagcctat gtgtgggaaa atcgtaatgg cacaaacacc    1620
aaccttgcag gcgcatggcc aggaacaaca gctacaaaag acgctgatgg attttatgtg    1680
tacaccataa caaatcacac acaatttagg gtaaatatta ttttcaacaa tgccagtggc    1740
acaaccgcca gcgagcaaac cgaagacctt tcggtaacgg ccgacatatg ttggcaggca    1800
gtaggcagca acaaatatgg tgtagaaaca gtgaactgcc caggcaccaa tgtcgaaaat    1860
aattacatac gtttattcaa catctatcca aatccgggca aagaaacagt gcgtataaac    1920
accaacgaaa cattgcgcgc agtaaccatg tacgatagca tgggacgcat gcaccgcatg    1980
```

caattaaatg ccgacaatag tttggacata agcaatatgg ccgccggact gtatcagatg 2040
caaattgaaa cagaaagcgg acgaaaagaa actcataaat ttgtaaaatt gtag 2094

<210> 363
<211> 697
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(20)

<400> 363
Met Lys Lys Ile Thr Phe Leu Leu Ala Leu Leu Leu Leu Val Ser Ala
1 5 10 15
Thr Ile Phe Ala Asn Asp Pro Lys Asp Asp Ile Met Leu Gln Ser Phe
20 25 30
Gly Trp Asp Glu Tyr Asp Gln Pro Lys Ile Lys Ala Ala Gly Asn Phe
35 40 45
Tyr Lys Phe Val Glu Ser Lys Ala Ser Glu Trp Lys Ala Ala Gly Phe
50 55 60
Asp Met Ile Trp Met Pro Pro Pro Ser Arg Ser Thr Gly Gly Val Gly
65 70 75 80
Tyr Leu Pro Thr Glu Leu His Asn Phe Asn Ser Thr Trp Gly Ser Glu
85 90 95
Ala Asp Leu Arg Ala Ala Thr Ala Ala Leu Lys Ala His Gly Met Phe
100 105 110
Pro Ile Ala Asp Val Val Ala Asn His Arg Asn Gly Thr Thr Ala Trp
115 120 125
Thr Asp Phe Thr Asn Pro Ala Trp Ser Cys Ala Thr Ile Thr Ser Thr
130 135 140
Asp Glu Ala Asn Val Asn Trp Val Ala Ser Arg Gly Pro Lys Pro Cys
145 150 155 160
Gly Ala Pro Asp Thr Gly Asp Asp Phe Asn Gly Gly Arg Asp Leu Asp
165 170 175
His Thr Asn Ile Glu Thr Arg Asn Gly Ile Lys Thr Phe Leu Thr Lys
180 185 190
Leu Lys Glu Gln Gly Phe Glu Gly Trp Arg Trp Asp Met Thr Lys Gly
195 200 205
Phe Ser Ala Ser Tyr Val Gly Glu Tyr Asn Thr Ala Ser Gln Pro Tyr
210 215 220
Phe Ser Val Gly Glu Tyr Trp Asp Gly Asn Ser Asn Thr Leu Lys Asn
225 230 235 240
Trp Val Asp Gly Thr Gly Lys Lys Ser Ala Thr Phe Asp Phe Ser Gln
245 250 255
Tyr Tyr Ser Met Glu Asn Ala Phe Lys Asn Asn Ala Trp Ser Ala Leu
260 265 270
Gly Ser Gly Ser Ser Met Ala Gly Leu Ala Gly Val Phe Gly Tyr Ser
275 280 285
Asp Tyr Ala Val Thr Phe Val Asp Asn His Asp Thr Phe Val His Gly
290 295 300
Ser Ala Pro Leu Gly Asp Asn Ile Met Lys Ala Tyr Ala Tyr Ile Leu
305 310 315 320
Thr His Pro Gly Ile Pro Ser Val Phe Ile Ala His Tyr Tyr Gly Gly
325 330 335
Thr Tyr Arg Lys Asp Gly Val Thr Arg Thr Tyr Thr Ser His Lys Asp
340 345 350
Lys Ile Asp Pro Ile Met Ala Val Arg Lys Ala Asn Arg Ile Asp Ala
355 360 365
Trp Ser Thr Leu Gln Val Ala Thr Asn Ser Gly Met Tyr Ala Ala Tyr
370 375 380
Ile Lys Arg Arg His Asp Asp Ala Ala Ala Ser Val Ala Val Lys Ile
385 390 395 400
Gly Pro Gly Ser Trp Ser Pro Glu Gly Ser Gly Trp Ile Leu Ala Ala
405 410 415
Ser Gly Thr Asp Tyr Ala Val Trp Ser Lys Ala Gln Ile Val Val Ala

```
            420                 425                 430
Pro Thr Met Leu Ile Gly Ser Gly Thr Tyr Glu Leu Gly Gln Thr Leu
        435                 440                 445
Ser Ile Thr Ala Pro Ser Gly Tyr Ser Ile Arg Tyr Thr Thr Asp Gly
    450                 455                 460
Ser Glu Pro Thr Ala Ala Ser Thr Leu Tyr Thr Ala Pro Ile Thr Leu
465                 470                 475                 480
Pro Gln Gly Thr Thr Val Ile Lys Ala Ala Ser Phe Val Asn Gly Val
                485                 490                 495
Met Ser Ala Val Val Thr Asn Thr Tyr Thr Ala Gln Ala Lys Ala Thr
            500                 505                 510
Ser Ile Lys Val Arg Phe Lys Ala Pro Ala Ser Trp Thr Ala Ser Lys
        515                 520                 525
Ala Tyr Val Trp Glu Asn Arg Asn Gly Thr Asn Thr Asn Leu Ala Gly
    530                 535                 540
Ala Trp Pro Gly Thr Thr Ala Thr Lys Asp Ala Asp Gly Phe Tyr Val
545                 550                 555                 560
Tyr Thr Ile Thr Asn His Thr Gln Phe Arg Val Asn Ile Ile Phe Asn
                565                 570                 575
Asn Ala Ser Gly Thr Thr Ala Ser Glu Gln Thr Glu Asp Leu Ser Val
            580                 585                 590
Thr Ala Asp Ile Cys Trp Gln Ala Val Gly Ser Asn Lys Tyr Gly Val
        595                 600                 605
Glu Thr Val Asn Cys Pro Gly Thr Asn Val Glu Asn Asn Tyr Ile Arg
    610                 615                 620
Leu Phe Asn Ile Tyr Pro Asn Pro Gly Lys Glu Thr Val Arg Ile Asn
625                 630                 635                 640
Thr Asn Glu Thr Leu Arg Ala Val Thr Met Tyr Asp Ser Met Gly Arg
                645                 650                 655
Met His Arg Met Gln Leu Asn Ala Asp Asn Ser Leu Asp Ile Ser Asn
            660                 665                 670
Met Ala Ala Gly Leu Tyr Gln Met Gln Ile Glu Thr Glu Ser Gly Arg
        675                 680                 685
Lys Glu Thr His Lys Phe Val Lys Leu
    690                 695
```

<210> 364
<211> 1536
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 364

```
atgaaaaatt ttaaaaagaa actaagtgca ttgttattat cctttattct tattttctct      60
agttttaatg ctgctatagt tcaagcagct accaatggca caatgtttca atactttgaa     120
tggtacttac caaatgacgg aacgctatgg gataaattag gtaatgatgc tccaaactta     180
aagagtatag gtgtaacagc tgtttggata ccaccagcat ataagggaag ttcacaggct     240
gatgttggat atggtaccta tgatttatat gacctaggtg aatttaatca aaaaggtaca     300
gtgagaacaa agtatggaac aaagagccaa ttaatcagtg cagtagataa actacataat     360
aatggaatac aagtatatgg agatgttgta ttaaatcata agatgggtgc tgatggaact     420
gaaacagttt gggctgttga agtaaatcca agcaatagaa atcaagaaac atcaggagat     480
tatcaaattt cagcatggac aaaatttgat tttccaggaa gaggaaacac atactcaagc     540
ttcaaatgga gatggtatca ttttgatgga gttgattggg accaaagcag aagcttaagt     600
agaatatata aattaagagg agactataaa ggatgggact gggaagttga tactgaaaat     660
ggaaattatg actatttaat gggtgcggat gtagacatgc aacaccctga agttgtaaat     720
gaattaaata actggggtaa gtggtttacc caaacagtaa aattagatgg atatagattg     780
gatgcagtaa agcatattaa gtttgatttc tttaatacat ggcttgatta tcaaagacaa     840
caatcaggaa aggaactatt tacagtaggt gaatactggt cagatgattt gggaaaactt     900
caaaattata taagtaaaac aggaagaaga atgtctctat ttgatgttcc actacaaaga     960
aatttctata atgcttcaaa tggtggtgga tattatgata tgagaaatat aaagagtgga    1020
acattaataa gtgttgaccc aactaaagca gttacatttg ttgacaacca tgatacagaa    1080
ccaggacaat cactacaaag atttgttcaa gattggttta agccacttgc atatacattt    1140
attttaacaa gagaagaagg atatccatgt atattctatg gtgattatta tggaattcca    1200
gcaaagggag tatcagcaaa gaaggtagca attgataaga taatgcttgc aagaaagaat    1260
tatgcttatg gaaagcaaaa tgactattta gaccactggg atataatagg ttggacaaga    1320
gaaggagata atgaacatgc taattcaggt atggcagcac ttataactga tggacctggt    1380
```

ggttcaaaat ggatgtatgt aggacaaaaa ttcataggaa agacattcta cgatattaca 1440
ggaaacagaa cagatacagt tgttataaat tcatcaggtt ggggtgaatt taaagtaaat 1500
ggtggttcat actcaatctg ggttccaaag aattaa 1536

<210> 365
<211> 511
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(29)

<400> 365

```
Met Lys Asn Phe Lys Lys Lys Leu Ser Ala Leu Leu Leu Ser Phe Ile
 1               5                  10                  15
Leu Ile Phe Ser Ser Phe Asn Ala Ala Ile Val Gln Ala Ala Thr Asn
            20                  25                  30
Gly Thr Met Phe Gln Tyr Phe Glu Trp Tyr Leu Pro Asn Asp Gly Thr
        35                  40                  45
Leu Trp Asp Lys Leu Gly Asn Asp Ala Pro Asn Leu Lys Ser Ile Gly
    50                  55                  60
Val Thr Ala Val Trp Ile Pro Pro Ala Tyr Lys Gly Ser Ser Gln Ala
65                  70                  75                  80
Asp Val Gly Tyr Gly Thr Tyr Asp Leu Tyr Asp Leu Gly Glu Phe Asn
                85                  90                  95
Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly Thr Lys Ser Gln Leu Ile
            100                 105                 110
Ser Ala Val Asp Lys Leu His Asn Asn Gly Ile Gln Val Tyr Gly Asp
        115                 120                 125
Val Val Leu Asn His Lys Met Gly Ala Asp Gly Thr Glu Thr Val Trp
    130                 135                 140
Ala Val Glu Val Asn Pro Ser Asn Arg Asn Gln Glu Thr Ser Gly Asp
145                 150                 155                 160
Tyr Gln Ile Ser Ala Trp Thr Lys Phe Asp Phe Pro Gly Arg Gly Asn
                165                 170                 175
Thr Tyr Ser Ser Phe Lys Trp Arg Trp Tyr His Phe Asp Gly Val Asp
            180                 185                 190
Trp Asp Gln Ser Arg Ser Leu Ser Arg Ile Tyr Lys Leu Arg Gly Asp
        195                 200                 205
Tyr Lys Gly Trp Asp Trp Glu Val Asp Thr Glu Asn Gly Asn Tyr Asp
    210                 215                 220
Tyr Leu Met Gly Ala Asp Val Asp Met Gln His Pro Glu Val Val Asn
225                 230                 235                 240
Glu Leu Asn Asn Trp Gly Lys Trp Phe Thr Gln Thr Val Lys Leu Asp
                245                 250                 255
Gly Tyr Arg Leu Asp Ala Val Lys His Ile Lys Phe Asp Phe Phe Asn
            260                 265                 270
Thr Trp Leu Asp Tyr Gln Arg Gln Gln Ser Gly Lys Glu Leu Phe Thr
        275                 280                 285
Val Gly Glu Tyr Trp Ser Asp Asp Leu Gly Lys Leu Gln Asn Tyr Ile
    290                 295                 300
Ser Lys Thr Gly Arg Arg Met Ser Leu Phe Asp Val Pro Leu Gln Arg
305                 310                 315                 320
Asn Phe Tyr Asn Ala Ser Asn Gly Gly Gly Tyr Tyr Asp Met Arg Asn
                325                 330                 335
Ile Lys Ser Gly Thr Leu Ile Ser Val Asp Pro Thr Lys Ala Val Thr
            340                 345                 350
Phe Val Asp Asn His Asp Thr Glu Pro Gly Gln Ser Leu Gln Arg Phe
        355                 360                 365
Val Gln Asp Trp Phe Lys Pro Leu Ala Tyr Thr Phe Ile Leu Thr Arg
    370                 375                 380
Glu Glu Gly Tyr Pro Cys Ile Phe Tyr Gly Asp Tyr Tyr Gly Ile Pro
385                 390                 395                 400
Ala Lys Gly Val Ser Ala Lys Lys Val Ala Ile Asp Lys Ile Met Leu
                405                 410                 415
```

```
Ala Arg Lys Asn Tyr Ala Tyr Gly Lys Gln Asn Asp Tyr Leu Asp His
            420                 425                 430
Trp Asp Ile Ile Gly Trp Thr Arg Glu Gly Asp Asn Glu His Ala Asn
        435                 440                 445
Ser Gly Met Ala Ala Leu Ile Thr Asp Gly Pro Gly Gly Ser Lys Trp
    450                 455                 460
Met Tyr Val Gly Gln Lys Phe Ile Gly Lys Thr Phe Tyr Asp Ile Thr
465                 470                 475                 480
Gly Asn Arg Thr Asp Thr Val Val Ile Asn Ser Ser Gly Trp Gly Glu
                485                 490                 495
Phe Lys Val Asn Gly Gly Ser Tyr Ser Ile Trp Val Pro Lys Asn
            500                 505                 510
```

<210> 366
<211> 1992
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 366

```
atgaaaatct ctagttttaa ctgcgcgaca cttggcgcag ggttattact ttcctccgca      60
caggtaagcg cagcgccaac gacctttgtg catttattcg attggaattg gcaagatatc     120
gctaccgagt gtgaaaccta tttaggtcca caaggttacg ctgccgtgca ggtttcgcct     180
ccaaatgaac atattgaagt gtctgagtgg tgggcgcgat atcagcctgt aagctacgtt     240
ctgcaaagcc gcagtggcag ccgcagtgag tttatcaata tggttgagcg ctgtgatgca     300
gtcggtgtag atatttacgt cgatgctgtc attaaccaca tggccgctgg aagcggtact     360
ggaactgcag gcaatacatt tggcaataag tcttatccta tctactctcc acaggacttt     420
catgctacat gcgctattaa tgactacggc gatcgctggc aagtacagaa ctgcgagtta     480
gtcgggttgg ccgacttaaa taccgccgat aattatgtac agacgacgct ggccggatat     540
ttaaatgatc ttgtgaacat tggtgttgcc ggtttccgtc tggacgcgtc aaaacacatg     600
gctgcaaccg atattgaagg tatcttacag agagtaaatg gcgatccgct tatttttcag     660
gaagtgattg atcaaggcgg cgaagcggta agtgcttcag agtattttgg gaacggcctg     720
gtaaccgaat ttaagtattc tgtagaaata ggaaatacat ttcgcaacgg ctcgctggca     780
tggctgagta acttcggcga agcgtgggga ttcatgccta gctatcaggc agtcgtgttt     840
gttgataatc acgacaatca gcgtggtcat ggtggaggtg gaaatgtcat caccttgag      900
gatggcagac tgtatgattt agccaatgta tttatgctgg cgtggccgta cggctaccct     960
aaagtaatgt ccagctttga ctttcatggc gatacagatg cgggacctcc tggcgttccg    1020
gtccataacg acggcaacct cgaatgcttt ggcagtaact ggaagtgtga acaccgctgg    1080
agctacattg caggcggtgt aaactttaga aatagcaccg cggacaattt cactgtcacc    1140
gactggtgga gcaacggtgc caaccaaatc gccttcggcc ggtctgacag tggctttgtt    1200
gtcattaata aagaatctta tgaccttaac accactctcg aaacgtcaat ggcaccggga    1260
gagtattgtg atgttctcac aggcgagctc aatgcatcag gcgatggctg taccggcaat    1320
aagctttcgg tagatagcaa tgggcgcctt tctgtctccg tcccatcctg ggaagcgatt    1380
gccattcacc ataatgccaa agtggcagcc acatcaaatt cggcagactg gcaacgaacg    1440
atgatattca tcgaagccga aactgccagt ggtcaagata tgtttgttcg gggcgacatt    1500
gaccatcaag ccgcagcgaa ccaactcggc atcacctgta cagatacaaa ttacgcttgt    1560
gcagtgccga ttgtgcataa taacttgcgc aatacgacta cggcttcctg gaaggcaaat    1620
gacaactatc ttgactggta tggtgttgaa gactcccaga gtactgacgc agaaggcagt    1680
gcgctggact ggacaaccaa ctcctggccc gcagaatggg gaaccttaaa aaccgtgatt    1740
gacgatggtt tcggtgaaac gccctataat acctggggct cgcatttctg gatgctcgac    1800
gtacaaatgg attgcagtaa gacagtgaat gggtggttcg aacttaaagc ctatgtcaaa    1860
aacgggcagg gctgggaagg cgatatcagt cagagcggta caccctaccc atcaaccaac    1920
cacatggcgc aatgcggtaa aaagaatatg ttccggtttg gacaaagcac ggtgcaaatc    1980
agcgatctgt aa                                                        1992
```

<210> 367
<211> 663
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(24)

<400> 367

```
Met Lys Ile Ser Ser Phe Asn Cys Ala Thr Leu Gly Ala Gly Leu Leu
 1               5                  10                  15
Leu Ser Ser Ala Gln Val Ser Ala Ala Pro Thr Thr Phe Val His Leu
            20                  25                  30
Phe Asp Trp Asn Trp Gln Asp Ile Ala Thr Glu Cys Glu Thr Tyr Leu
        35                  40                  45
Gly Pro Gln Gly Tyr Ala Ala Val Gln Val Ser Pro Pro Asn Glu His
    50                  55                  60
Ile Glu Val Ser Glu Trp Trp Ala Arg Tyr Gln Pro Val Ser Tyr Val
65                  70                  75                  80
Leu Gln Ser Arg Ser Gly Ser Arg Ser Glu Phe Ile Asn Met Val Glu
                85                  90                  95
Arg Cys Asp Ala Val Gly Val Asp Ile Tyr Val Asp Ala Val Ile Asn
            100                 105                 110
His Met Ala Ala Gly Ser Gly Thr Gly Thr Ala Gly Asn Thr Phe Gly
        115                 120                 125
Asn Lys Ser Tyr Pro Ile Tyr Ser Pro Gln Asp Phe His Ala Thr Cys
    130                 135                 140
Ala Ile Asn Asp Tyr Gly Asp Arg Trp Gln Val Gln Asn Cys Glu Leu
145                 150                 155                 160
Val Gly Leu Ala Asp Leu Asn Thr Ala Asp Asn Tyr Val Gln Thr Thr
                165                 170                 175
Leu Ala Gly Tyr Leu Asn Asp Leu Val Asn Ile Gly Val Ala Gly Phe
            180                 185                 190
Arg Leu Asp Ala Ser Lys His Met Ala Ala Thr Asp Ile Glu Gly Ile
        195                 200                 205
Leu Gln Arg Val Asn Gly Asp Pro Leu Ile Phe Gln Glu Val Ile Asp
    210                 215                 220
Gln Gly Gly Glu Ala Val Ser Ala Ser Glu Tyr Phe Gly Asn Gly Leu
225                 230                 235                 240
Val Thr Glu Phe Lys Tyr Ser Val Glu Ile Gly Asn Thr Phe Arg Asn
                245                 250                 255
Gly Ser Leu Ala Trp Leu Ser Asn Phe Gly Glu Ala Trp Gly Phe Met
            260                 265                 270
Pro Ser Tyr Gln Ala Val Val Phe Val Asp Asn His Asp Asn Gln Arg
        275                 280                 285
Gly His Gly Gly Gly Gly Asn Val Ile Thr Phe Glu Asp Gly Arg Leu
    290                 295                 300
Tyr Asp Leu Ala Asn Val Phe Met Leu Ala Trp Pro Tyr Gly Tyr Pro
305                 310                 315                 320
Lys Val Met Ser Ser Phe Asp Phe His Gly Asp Thr Asp Ala Gly Pro
                325                 330                 335
Pro Gly Val Pro Val His Asn Asp Gly Asn Leu Glu Cys Phe Gly Ser
            340                 345                 350
Asn Trp Lys Cys Glu His Arg Trp Ser Tyr Ile Ala Gly Gly Val Asn
        355                 360                 365
Phe Arg Asn Ser Thr Ala Asp Asn Phe Thr Val Thr Asp Trp Trp Ser
    370                 375                 380
Asn Gly Ala Asn Gln Ile Ala Phe Gly Arg Ser Asp Ser Gly Phe Val
385                 390                 395                 400
Val Ile Asn Lys Glu Ser Tyr Asp Leu Asn Thr Thr Leu Glu Thr Ser
                405                 410                 415
Met Ala Pro Gly Glu Tyr Cys Asp Val Leu Thr Gly Glu Leu Asn Ala
            420                 425                 430
Ser Gly Asp Gly Cys Thr Gly Asn Lys Leu Ser Val Asp Ser Asn Gly
        435                 440                 445
Arg Leu Ser Val Ser Val Pro Ser Trp Glu Ala Ile Ala Ile His His
    450                 455                 460
Asn Ala Lys Val Ala Ala Thr Ser Asn Ser Ala Asp Trp Gln Arg Thr
465                 470                 475                 480
Met Ile Phe Ile Glu Ala Glu Thr Ala Ser Gly Gln Asp Met Phe Val
                485                 490                 495
Arg Gly Asp Ile Asp His Gln Ala Ala Ala Asn Gln Leu Gly Ile Thr
            500                 505                 510
Cys Thr Asp Thr Asn Tyr Ala Cys Ala Val Pro Ile Val His Asn Asn
        515                 520                 525
Leu Arg Asn Thr Thr Thr Ala Ser Trp Lys Ala Asn Asp Asn Tyr Leu
```

```
    530                 535                 540
Asp Trp Tyr Gly Val Glu Asp Ser Gln Ser Thr Asp Ala Glu Gly Ser
545                 550                 555                 560
Ala Leu Asp Trp Thr Thr Asn Ser Trp Pro Ala Glu Trp Gly Thr Leu
                565                 570                 575
Lys Thr Val Ile Asp Asp Gly Phe Gly Glu Thr Pro Tyr Asn Thr Trp
            580                 585                 590
Gly Ser His Phe Trp Met Leu Asp Val Gln Met Asp Cys Ser Lys Thr
        595                 600                 605
Val Asn Gly Trp Phe Glu Leu Lys Ala Tyr Val Lys Asn Gly Gln Gly
    610                 615                 620
Trp Glu Gly Asp Ile Ser Gln Ser Gly Thr Pro Tyr Pro Ser Thr Asn
625                 630                 635                 640
His Met Ala Gln Cys Gly Lys Lys Asn Met Phe Arg Phe Gly Gln Ser
                645                 650                 655
Thr Val Gln Ile Ser Asp Leu
            660
```

<210> 368
<211> 1257
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 368

```
atgtttattt acaatctttt ccctaggctt tatgggccat ttagcaagtg gatagacgac      60
ttagatagaa taaaggagtt gggagtagac tggatatata taaaccccgt atcttatccc     120
gggttttccg gcagcctata ctcaataaaa gatcattaca aattcaaccc tctgttccta     180
gactctaatt caaaaaaatc accagaggaa caattcagag atttcattaa agcgtgtaaa     240
tcaagaggaa taaaagtaat cgttgatttg gtagtcaatc acacagctat agattccgta     300
cttgtagaca aacatcctaa ctggtacaaa agagacgaaa acgggcaaat aaagagacca     360
ggcgcatggg ataatagcgt atgggttgaa tggggagatt tagcagaact agacaacgag     420
aacagtcctg ataaaagagc attatggaac tattggaaag acctagtaga gtggcatata     480
gacttgggag ttgatgggtt tagatgcgac tatgcttata atgtccctgt agaattatgg     540
gacttcctaa taaagaacgc aaaatccaag aaagaagatg ttatattctt agcagagata     600
ctcggaggtc cctttgatag aaacctagaa gttgcaaaag ctggctttga ttatgtattt     660
agcagtgcta aatggtggaa tttctcagac gcttggttca tagaacaaac aagagagttc     720
tcaaaatatg ttaaagcaat agcatttcct gaatctcatg acacaccaag gttgaattcg     780
gagtacgaag aagatttagc aaaaataaag caaagattca tatttacagc attcataaat     840
gaaggattaa tgatacctgt aggatttgag tttggcttca ttaacaagct agatgtggta     900
aaaacaaccc cagactggtg ggaaatacca tcttatgaca tttctaagtt cataaaacac     960
atagttgaaa tcaaaaaaaa gcatcccata ttgtcatatg acgctacatc agtagatatg    1020
ctagaaaaca cccctgacat ctcaatgttt gtcaagagga acaagctaac agaccagatt    1080
gccctttcg tagttaacaa aaacacaacc tcaaaagtaa agagaataat agatttccct    1140
agtgtattcg gttattatga tgctatagat gtaatagata acgttgtaat gcctaagtac    1200
gctgaagtaa cagtaccgag agcaggtata aaagttttcg taaacagaat acactaa       1257
```

<210> 369
<211> 418
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 369

```
Met Phe Ile Tyr Asn Leu Phe Pro Arg Leu Tyr Gly Pro Phe Ser Lys
 1               5                  10                  15
Trp Ile Asp Asp Leu Asp Arg Ile Lys Glu Leu Gly Val Asp Trp Ile
            20                  25                  30
Tyr Ile Asn Pro Val Ser Tyr Pro Gly Phe Ser Gly Ser Leu Tyr Ser
        35                  40                  45
Ile Lys Asp His Tyr Lys Phe Asn Pro Leu Phe Leu Asp Ser Asn Ser
    50                  55                  60
Lys Lys Ser Pro Glu Glu Gln Phe Arg Asp Phe Ile Lys Ala Cys Lys
65                  70                  75                  80
```

```
Ser Arg Gly Ile Lys Val Ile Val Asp Leu Val Val Asn His Thr Ala
                85                  90                  95
Ile Asp Ser Val Leu Val Asp Lys His Pro Asn Trp Tyr Lys Arg Asp
            100                 105                 110
Glu Asn Gly Gln Ile Lys Arg Pro Gly Ala Trp Asp Asn Ser Val Trp
        115                 120                 125
Val Glu Trp Gly Asp Leu Ala Glu Leu Asp Asn Glu Asn Ser Pro Asp
    130                 135                 140
Lys Arg Ala Leu Trp Asn Tyr Trp Lys Asp Leu Val Glu Trp His Ile
145                 150                 155                 160
Asp Leu Gly Val Asp Gly Phe Arg Cys Asp Tyr Ala Tyr Asn Val Pro
                165                 170                 175
Val Glu Leu Trp Asp Phe Leu Ile Lys Asn Ala Lys Ser Lys Lys Glu
            180                 185                 190
Asp Val Ile Phe Leu Ala Glu Ile Leu Gly Gly Pro Phe Asp Arg Asn
        195                 200                 205
Leu Glu Val Ala Lys Ala Gly Phe Asp Tyr Val Phe Ser Ser Ala Lys
    210                 215                 220
Trp Trp Asn Phe Ser Asp Ala Trp Phe Ile Glu Gln Thr Arg Glu Phe
225                 230                 235                 240
Ser Lys Tyr Val Lys Ala Ile Ala Phe Pro Glu Ser His Asp Thr Pro
                245                 250                 255
Arg Leu Asn Ser Glu Tyr Glu Glu Asp Leu Ala Lys Ile Lys Gln Arg
            260                 265                 270
Phe Ile Phe Thr Ala Phe Ile Asn Glu Gly Leu Met Ile Pro Val Gly
        275                 280                 285
Phe Glu Phe Gly Phe Ile Asn Lys Leu Asp Val Val Lys Thr Thr Pro
    290                 295                 300
Asp Trp Trp Glu Ile Pro Ser Tyr Asp Ile Ser Lys Phe Ile Lys His
305                 310                 315                 320
Ile Val Glu Ile Lys Lys Lys His Pro Ile Leu Ser Tyr Asp Ala Thr
                325                 330                 335
Ser Val Asp Met Leu Glu Asn Thr Pro Asp Ile Ser Met Phe Val Lys
            340                 345                 350
Arg Asn Lys Leu Thr Asp Gln Ile Ala Leu Phe Val Val Asn Lys Asn
        355                 360                 365
Thr Thr Ser Lys Val Lys Arg Ile Ile Asp Phe Pro Ser Val Phe Gly
    370                 375                 380
Tyr Tyr Asp Ala Ile Asp Val Ile Asp Asn Val Val Met Pro Lys Tyr
385                 390                 395                 400
Ala Glu Val Thr Val Pro Arg Ala Gly Ile Lys Val Phe Val Asn Arg
                405                 410                 415
Ile His


<210> 370
<211> 1614
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 370
atgaaacgct acttaacact caccctactg cttaccctgt tgatacttgc gccgctggct      60
ttggcacaag agtccgatga gggcgcagca tgggctaacg agactgtttt ttatgagatt     120
tttgtgcgta gcttttacga tagcgacggt gacggtaaag gggatattcg cggcgttatc     180
gaaaaactcg attacctcaa cgacggcgac ccgaatacca caaccgactt aggcattacc     240
ggcatttggt tgatgcccgt catggagtcg cccagttatc acggctacga cgcgaccgat     300
tactacatgc tcgataaaga ctacggcaca aacgacgact tcaagctgtt catggaagaa     360
gcgcacaaac gcggaatccg tgtcatcgtt gacttgatgc tcaatcacac ctcggtgcag     420
catgagtggt tcatcgaatc gggtgacccg caaagcgaaa aagccgattg gtacatctgg     480
gaagatgaga acccgaatta tcgcggcccg gataatcagc aggtgtggca cggcctacgc     540
gggcgctatt actacggcgt gttctggtca gggatgccgg acttgaactt caccaatccc     600
gatgtcaccg cgcaaatgta cgatgtcacg cgcttttggc tagaagagat gaacgtggac     660
ggcttccgcc ttgacgcgat taaacacatc atcgaggatg gcaccattca agagaacacc     720
cccgccacgc gccaatggct ggccgattac cgcgcgttcg tcaagtccat caagcctgac     780
gcgctgttgg tcggtgaggt ctggtcgccg acgttcgcta tcgcgccgta tgtgggcaca     840
```

gcggttgaca tcgcgtttga gttcgacttg gccgaggcgc tgatacgcgg cgcgagcttt 900
ggcacgccca acctcgtcac gcgacagcta gagacggtct ttgagagtta cccgcccaac 960
gctttcgcca cattcttgac taaccacgat cagaaccgct tcatgaccca aattcgcggg 1020
gacgcggcag cggccaaaaa cggcgcatcg ttgctactca cgttgcccgg ggtaccgttc 1080
atttactacg gtgaagaaat cggcatgacc ggcatgaaac ctgatgagga catccgcaaa 1140
ccgatgcagt gggacgatac ccccaagacc gctggtttca ccagtggtga gcgcgtgtgg 1200
cgtgctgtga ataacgatgc attcgagggc gtcaatgtgg ccgcacaaac cgaaaacccc 1260
gactcacttc taagccacta ccgcgcgctg gttcaagcac gggtgagcag ccccgcgtta 1320
caaagcggcg acatgatttt ggttgaatcg tctaaccgta agctgtttac tttcttgcgc 1380
caaagcgagg cccaaacggt gctggtgatg attaaccttg acgaccgtcc cgtgaccgag 1440
tacacgctca cgctggaaag tggcacactc gcggaaggcg tgacagcatc ggcgctcctc 1500
gcatcaggtg aggtaaacgc gccacaggtc aacgggcaag gcggctttgc ggattacacc 1560
ccgcttgcag aactcgcgcc caattcggtc acgctgattc agcttgggga ataa 1614

<210> 371
<211> 537
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(22)

<400> 371
Met Lys Arg Tyr Leu Thr Leu Thr Leu Leu Leu Thr Leu Leu Ile Leu
 1   5   10   15
Ala Pro Leu Ala Leu Ala Gln Glu Ser Asp Glu Gly Ala Ala Trp Ala
            20   25   30
Asn Glu Thr Val Phe Tyr Glu Ile Phe Val Arg Ser Phe Tyr Asp Ser
        35   40   45
Asp Gly Asp Gly Lys Gly Asp Ile Arg Gly Val Ile Glu Lys Leu Asp
    50   55   60
Tyr Leu Asn Asp Gly Asp Pro Asn Thr Thr Thr Asp Leu Gly Ile Thr
65   70   75   80
Gly Ile Trp Leu Met Pro Val Met Glu Ser Pro Ser Tyr His Gly Tyr
                85   90   95
Asp Ala Thr Asp Tyr Tyr Met Leu Asp Lys Asp Tyr Gly Thr Asn Asp
            100   105   110
Asp Phe Lys Leu Phe Met Glu Glu Ala His Lys Arg Gly Ile Arg Val
        115   120   125
Ile Val Asp Leu Met Leu Asn His Thr Ser Val Gln His Glu Trp Phe
    130   135   140
Ile Glu Ser Gly Asp Pro Gln Ser Glu Lys Ala Asp Trp Tyr Ile Trp
145   150   155   160
Glu Asp Glu Asn Pro Asn Tyr Arg Gly Pro Asp Asn Gln Gln Val Trp
                165   170   175
His Gly Leu Arg Gly Arg Tyr Tyr Tyr Gly Val Phe Trp Ser Gly Met
            180   185   190
Pro Asp Leu Asn Phe Thr Asn Pro Asp Val Thr Ala Gln Met Tyr Asp
        195   200   205
Val Thr Arg Phe Trp Leu Glu Glu Met Asn Val Asp Gly Phe Arg Leu
    210   215   220
Asp Ala Ile Lys His Ile Ile Glu Asp Gly Thr Ile Gln Glu Asn Thr
225   230   235   240
Pro Ala Thr Arg Gln Trp Leu Ala Asp Tyr Arg Ala Phe Val Lys Ser
                245   250   255
Ile Lys Pro Asp Ala Leu Leu Val Gly Glu Val Trp Ser Pro Thr Phe
            260   265   270
Ala Ile Ala Pro Tyr Val Gly Thr Ala Val Asp Ile Ala Phe Glu Phe
        275   280   285
Asp Leu Ala Glu Ala Leu Ile Arg Gly Ala Ser Phe Gly Thr Pro Asn
    290   295   300
Leu Val Thr Arg Gln Leu Glu Thr Val Phe Glu Ser Tyr Pro Pro Asn
305   310   315   320
Ala Phe Ala Thr Phe Leu Thr Asn His Asp Gln Asn Arg Phe Met Thr
                325   330   335

```
Gln Ile Arg Gly Asp Ala Ala Ala Ala Lys Asn Gly Ala Ser Leu Leu
            340                 345                 350
Leu Thr Leu Pro Gly Val Pro Phe Ile Tyr Tyr Gly Glu Glu Ile Gly
        355                 360                 365
Met Thr Gly Met Lys Pro Asp Glu Asp Ile Arg Lys Pro Met Gln Trp
    370                 375                 380
Asp Asp Thr Pro Lys Thr Ala Gly Phe Thr Ser Gly Glu Arg Val Trp
385                 390                 395                 400
Arg Ala Val Asn Asn Asp Ala Phe Glu Gly Val Asn Val Ala Ala Gln
                405                 410                 415
Thr Glu Asn Pro Asp Ser Leu Leu Ser His Tyr Arg Ala Leu Val Gln
            420                 425                 430
Ala Arg Val Ser Ser Pro Ala Leu Gln Ser Gly Asp Met Ile Leu Val
        435                 440                 445
Glu Ser Ser Asn Arg Lys Leu Phe Thr Phe Leu Arg Gln Ser Glu Ala
    450                 455                 460
Gln Thr Val Leu Val Met Ile Asn Leu Asp Asp Arg Pro Val Thr Glu
465                 470                 475                 480
Tyr Thr Leu Thr Leu Glu Ser Gly Thr Leu Ala Glu Gly Val Thr Ala
                485                 490                 495
Ser Ala Leu Leu Ala Ser Gly Glu Val Asn Ala Pro Gln Val Asn Gly
            500                 505                 510
Gln Gly Gly Phe Ala Asp Tyr Thr Pro Leu Ala Glu Leu Ala Pro Asn
        515                 520                 525
Ser Val Thr Leu Ile Gln Leu Gly Glu
    530                 535
```

<210> 372
<211> 1437
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 372
atgttcaaaa tagtgccgct cctttctctc gtcacggctt tcgtcatcgg cgcatacccg 60
ctgcgcgcgc agtccgctat agaggacctg gtaaaggcct ctccggtccg ggaccaggcc 120
gtaccggccc cctccgagcc caaatccgcc cacacgccca gcgccctggc gcgccaggac 180
tggaaaacgg ccgtaaccct ctattcggcc aagaacttca atcctccgcc cgcgcggcgg 240
ctgaacctcg aggcctcgcc ggtggtcatg ctgcagggct tccactggta cgccgacaat 300
tactggctca acctgccccg cggctggtgg ggcaggctgg ccgagaaagc cccggaggtc 360
ggcgccgccg gcttcgacct ggtctggttc cccccggtat cccgcggcag ctattacccc 420
accgagtggt acgacctcga ttcccagtgg ggcaagaaag acgtcctgat agaggccgtg 480
cgggcgatga aggcccacgg ggtcagcccg gtggccgaca tcatactcaa ccaccgcaac 540
ggcaccctcg actgggcgga tttcaccagc cccgactggc cgaccacggc gatagtcagg 600
gacgacgagt ggcagggcat gaaaggcccc aattacgacg agggacaggg agagggcggc 660
tgccgcgacc tggatcacaa caacgccctg gtgcgcgagg acgccaaggt cttcctgcgc 720
tggctgcgct acaccgtcgg tttcgaaggc tggcgctacg acatggtcaa gggcttcctg 780
ccccacagga tcaaggacta caaccgcgct tcggcgccgg ccttctcggt aggggaatac 840
tacgacacca accgccagct gctggtcaac tgggtggacg gcacggccga cggcgacaag 900
tcggaagcct ccacggtctt cgatttcacc acccgctaca acctggtggc cgccgtagag 960
agcgagcgtt acgacctgct cagcgaccac ggccgcccct ccggctccat cggctggtgg 1020
ccggccaagt cggtcacctt cgtggagaac cacgacacct cgccccgcga ccccaagttc 1080
ctgcagaacg ccccgcgcga gtaccgcgtc cagcgccaga tgggctacgc ctacatcctt 1140
acccatcccg gcacgcccag cgtcttctgg ccccatttct tcgactgggg ttccgaatac 1200
agggccaagc tgcaggagct cgtcaatatc cgcaagtcgg ccggcataac cgccgtcagc 1260
ccggtgcaga tactcgcggc ctacaacgag ctctacgccg cgatgataac cggcagcgat 1320
aagtacgtcg tattgaagct cggcaagaac tggggctggc agcccggcca cgactggaag 1380
ctggaagcca gcggcgagcg ctgggccgtc tggacccagc ccgtcaccag gccctga 1437

<210> 373
<211> 478
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(25)

<400> 373

```
Met Phe Lys Ile Val Pro Leu Leu Ser Leu Val Thr Ala Phe Val Ile
 1               5                  10                  15
Gly Ala Tyr Pro Leu Arg Ala Gln Ser Ala Ile Glu Asp Leu Val Lys
            20                  25                  30
Ala Ser Pro Val Arg Asp Gln Ala Val Pro Ala Pro Ser Glu Pro Lys
        35                  40                  45
Ser Ala His Thr Pro Ser Ala Leu Ala Arg Gln Asp Trp Lys Thr Ala
    50                  55                  60
Val Thr Leu Tyr Ser Ala Lys Asn Phe Asn Pro Pro Pro Ala Arg Arg
65                  70                  75                  80
Leu Asn Leu Glu Ala Ser Pro Val Val Met Leu Gln Gly Phe His Trp
                85                  90                  95
Tyr Ala Asp Asn Tyr Trp Leu Asn Leu Pro Arg Gly Trp Trp Gly Arg
            100                 105                 110
Leu Ala Glu Lys Ala Pro Glu Val Gly Ala Ala Gly Phe Asp Leu Val
        115                 120                 125
Trp Phe Pro Pro Val Ser Arg Gly Ser Tyr Tyr Pro Thr Glu Trp Tyr
    130                 135                 140
Asp Leu Asp Ser Gln Trp Gly Lys Lys Asp Val Leu Ile Glu Ala Val
145                 150                 155                 160
Arg Ala Met Lys Ala His Gly Val Ser Pro Val Ala Asp Ile Ile Leu
                165                 170                 175
Asn His Arg Asn Gly Thr Leu Asp Trp Ala Asp Phe Thr Ser Pro Asp
            180                 185                 190
Trp Pro Thr Thr Ala Ile Val Arg Asp Asp Glu Trp Gln Gly Met Lys
        195                 200                 205
Gly Pro Asn Tyr Asp Glu Gly Gln Gly Glu Gly Gly Cys Arg Asp Leu
    210                 215                 220
Asp His Asn Asn Ala Leu Val Arg Glu Asp Ala Lys Val Phe Leu Arg
225                 230                 235                 240
Trp Leu Arg Tyr Thr Val Gly Phe Glu Gly Trp Arg Tyr Asp Met Val
                245                 250                 255
Lys Gly Phe Leu Pro His Arg Ile Lys Asp Tyr Asn Arg Ala Ser Ala
            260                 265                 270
Pro Ala Phe Ser Val Gly Glu Tyr Tyr Asp Thr Asn Arg Gln Leu Leu
        275                 280                 285
Val Asn Trp Val Asp Gly Thr Ala Asp Gly Asp Lys Ser Glu Ala Ser
    290                 295                 300
Thr Val Phe Asp Phe Thr Thr Arg Tyr Asn Leu Val Ala Ala Val Glu
305                 310                 315                 320
Ser Glu Arg Tyr Asp Leu Leu Ser Asp His Gly Arg Pro Ser Gly Ser
                325                 330                 335
Ile Gly Trp Trp Pro Ala Lys Ser Val Thr Phe Val Glu Asn His Asp
            340                 345                 350
Thr Ser Pro Arg Asp Pro Lys Phe Leu Gln Asn Ala Pro Arg Glu Tyr
        355                 360                 365
Arg Val Gln Arg Gln Met Gly Tyr Ala Tyr Ile Leu Thr His Pro Gly
    370                 375                 380
Thr Pro Ser Val Phe Trp Pro His Phe Phe Asp Trp Gly Ser Glu Tyr
385                 390                 395                 400
Arg Ala Lys Leu Gln Glu Leu Val Asn Ile Arg Lys Ser Ala Gly Ile
                405                 410                 415
Thr Ala Val Ser Pro Val Gln Ile Leu Ala Ala Tyr Asn Glu Leu Tyr
            420                 425                 430
Ala Ala Met Ile Thr Gly Ser Asp Lys Tyr Val Val Leu Lys Leu Gly
        435                 440                 445
Lys Asn Trp Gly Trp Gln Pro Gly His Asp Trp Lys Leu Glu Ala Ser
    450                 455                 460
Gly Glu Arg Trp Ala Val Trp Thr Gln Pro Val Thr Arg Pro
465                 470                 475
```

<210> 374
<211> 1398

<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 374
atgaaacaca cagcgggaat atgggcgatc gcaggcatgc tgatcgcccc cttggcgcac 60
gccgacgtca tactgcacgc cttcaactgg aagtacagtg aagtgaccgc caaagccgat 120
cttatcaagt cagccggtta caaacaggta ctcatctcac cacctctcaa gtcatcaggc 180
aacgagtggt gggctcgtta ccagccgcag gatctgcgtc ttgtcgactc ccctcttggc 240
aacaagcagg atctcgagca gctgatcgcc gccatgaagg cgcgaggtat cgccgtctat 300
gccgatgtgg tgctcaacca catggccaac gaaagctgga aacgcagcga tctcaactat 360
cccggcaacg agctgctgca aagctacgcc agcaacccgg cctatttcga gcgccagaaa 420
ctgtttggcg atctggggca aaacgtcctc tccggtcagg attttcatcc cgaagggtgc 480
atcagcgact ggagcgatcc gggtcatgtc cagtattggc gactctgtgg tggtgccggt 540
gacaaggggc tgccggatct ggatcccaac aactgggtgg tcagccagca acaagcctat 600
ctcaaggcgc tcaaggggat ggggatcaag ggttttcggg tcgatgcggt caaacacatg 660
agcgattacc agatcaatgc tgtgttcacc cccgagatca aacaggggat gcacgtgttt 720
ggcgaggtga tcaccacagg aggggcaggc agcaccgatt acgagcgctt cctcaaaccc 780
tacctcgaca gcagcggcca gggagcctat gactttccgc tgtttgcctc cctgcgtggt 840
gcgttaagtt acggtggcag catgaacctg ctggcggatc cgggcgccta tggtcaggcg 900
ctaccgggca atcgcgccgt cacctttgcc gtgacccacg acatccccac caacgacggc 960
ttccgctatc agatcctgag ccagactgac gagaagctgg cctatgccta cctgctcggc 1020
cgcgatggtg gctcaccact ggtctattcg gatcacggcg agacccaggc caaggatgga 1080
ctgcgctggc aggattacta ccagcgcgcc gatctcaagg ggatgatccg tttccacaat 1140
gcggtgcagg gccagcccat gcagctggtc ggcagtggcg actgcttcgt gctgttccaa 1200
cggggcaaac aggggctggt cggcatcaac aagtgcgact atgagcagga gtactggctc 1260
gataccgcca ggttcgagat gaactggtat cgcaactatc gggacgtgct cgaccaaggg 1320
gccgtagtca acgtgcagag ccagtgggta cgggtcgcca ttcctgcccg cagcgccaga 1380
atgtggctgc aggagtga 1398

<210> 375
<211> 465
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(21)

<400> 375
Met Lys His Thr Ala Gly Ile Trp Ala Ile Ala Gly Met Leu Ile Ala
 1   5   10   15
Pro Leu Ala His Ala Asp Val Ile Leu His Ala Phe Asn Trp Lys Tyr
            20   25   30
Ser Glu Val Thr Ala Lys Ala Asp Leu Ile Lys Ser Ala Gly Tyr Lys
        35   40   45
Gln Val Leu Ile Ser Pro Pro Leu Lys Ser Ser Gly Asn Glu Trp Trp
    50   55   60
Ala Arg Tyr Gln Pro Gln Asp Leu Arg Leu Val Asp Ser Pro Leu Gly
65   70   75   80
Asn Lys Gln Asp Leu Glu Gln Leu Ile Ala Ala Met Lys Ala Arg Gly
                85   90   95
Ile Ala Val Tyr Ala Asp Val Val Leu Asn His Met Ala Asn Glu Ser
            100   105   110
Trp Lys Arg Ser Asp Leu Asn Tyr Pro Gly Asn Glu Leu Leu Gln Ser
        115   120   125
Tyr Ala Ser Asn Pro Ala Tyr Phe Glu Arg Gln Lys Leu Phe Gly Asp
    130   135   140
Leu Gly Gln Asn Val Leu Ser Gly Gln Asp Phe His Pro Glu Gly Cys
145   150   155   160
Ile Ser Asp Trp Ser Asp Pro Gly His Val Gln Tyr Trp Arg Leu Cys
                165   170   175
Gly Gly Ala Gly Asp Lys Gly Leu Pro Asp Leu Asp Pro Asn Asn Trp
            180   185   190

```
Val Val Ser Gln Gln Gln Ala Tyr Leu Lys Ala Leu Lys Gly Met Gly
        195                 200                 205
Ile Lys Gly Phe Arg Val Asp Ala Val Lys His Met Ser Asp Tyr Gln
    210                 215                 220
Ile Asn Ala Val Phe Thr Pro Glu Ile Lys Gln Gly Met His Val Phe
225                 230                 235                 240
Gly Glu Val Ile Thr Thr Gly Gly Ala Gly Ser Thr Asp Tyr Glu Arg
                245                 250                 255
Phe Leu Lys Pro Tyr Leu Asp Ser Ser Gly Gln Gly Ala Tyr Asp Phe
            260                 265                 270
Pro Leu Phe Ala Ser Leu Arg Gly Ala Leu Ser Tyr Gly Gly Ser Met
        275                 280                 285
Asn Leu Leu Ala Asp Pro Gly Ala Tyr Gly Gln Ala Leu Pro Gly Asn
    290                 295                 300
Arg Ala Val Thr Phe Ala Val Thr His Asp Ile Pro Thr Asn Asp Gly
305                 310                 315                 320
Phe Arg Tyr Gln Ile Leu Ser Gln Thr Asp Glu Lys Leu Ala Tyr Ala
                325                 330                 335
Tyr Leu Leu Gly Arg Asp Gly Gly Ser Pro Leu Val Tyr Ser Asp His
            340                 345                 350
Gly Glu Thr Gln Ala Lys Asp Gly Leu Arg Trp Gln Asp Tyr Tyr Gln
        355                 360                 365
Arg Ala Asp Leu Lys Gly Met Ile Arg Phe His Asn Ala Val Gln Gly
    370                 375                 380
Gln Pro Met Gln Leu Val Gly Ser Gly Asp Cys Phe Val Leu Phe Gln
385                 390                 395                 400
Arg Gly Lys Gln Gly Leu Val Gly Ile Asn Lys Cys Asp Tyr Glu Gln
                405                 410                 415
Glu Tyr Trp Leu Asp Thr Ala Arg Phe Glu Met Asn Trp Tyr Arg Asn
            420                 425                 430
Tyr Arg Asp Val Leu Asp Gln Gly Ala Val Val Asn Val Gln Ser Gln
        435                 440                 445
Trp Val Arg Val Ala Ile Pro Ala Arg Ser Ala Arg Met Trp Leu Gln
    450                 455                 460
Glu
465
```

<210> 376
<211> 1551
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 376

```
atgtcacaat tcctaaggtt cgcagtcagg cgcatatgta ctcagtggtg gtgccgtacc      60
actgggatta ttcttctggt ggttttcgcg ggcgtgtcgg cggcgttggg ccaggccggt     120
ttcgacgatg accgtgtcat gctgcaggga ttttattggg aatcctatcg gcatggacac     180
ctggagaagt tccccaatta cggcgataaa cactggtacg tgattgtaca aaacctcgca     240
ccggaaatcc gagccggccg cttcaatctg atctggctgc cgccgccctc gtacgcgggc     300
ggctttagcg ccggctataa tccaaaggag tactttcgcc ttgataacag ctacggcacg     360
ttcgatgagc accgcgccgc gcttgaagcg ctccttcgga acggaatcga gccagtggct     420
gatctagtca tcaatcaccg cgacggtaac caggcgtggg cggatttcaa aaatcccgat     480
tggggaactt gggcgatctg ccgaacggac gaggcattca caaatgtgaa ctccggcatc     540
accaatacgc cgaacgatca acgcggtaac tgtgaggaag aagccgatta tcggcccgag     600
ggcacgtata attattcgga ttttcgcgac attgcgcatg ctgatccgcg ggtgcggcgc     660
gacattgaac gatatctgct ccagttaaaa tctttgggtt accgggggtg gcgttacgat     720
atggtgcacg gctataacgc gagatgggtt gcgctctata atcgactcac gcaacccact     780
ttctcggtag gagagtacga ttgggacaag cacgcgcagc agcgtggctg gatatgggga     840
accgcaacga acccgactcc aacgggtgca gatcacctca gaacttcgag cagcgtcttt     900
gatttcacga gccagttcag tctgaaatca atcaacagcg gaaactacac ggcactgtat     960
ggcttcggaa acggcatcgg tctggtgggc gatacaacag atgggctgcc gtggaagaat    1020
cgcgccgtga cgtttctcga aaaccacgat accgggtacc gcaccaacga ggacggcacg    1080
ccgcagcagg gccaccaatt cgattcgttt gcgaacaact ggcaggtcga gcaggggtat    1140
gcgcagattt tgacacatcc cggtgtgccc actgtctact ggaagcatta ctttgactgg    1200
ggatcggacc tgcagaataa gatcagagca ctgatcaatg ctcggaaggt agcgggggtt    1260
aacgcgggca gtaaggtaga cccacaggac aatgccaaat actcgggcgt atacgcggct    1320
```

agaatcagcg gacgtaatgg cgagttgtac gtgcgcatag gagggagcga ccagcaatgg 1380
cagcctagtg cttcgggata tgccgattac cgcgagtacg cgcaaggagc cggctggaaa 1440
gtctgggtta aattacctgg aaatcccccg gtgcagcagg tgcctcttcc cagcgcgttg 1500
cctgttccgc agtaccggcc tccggaccaa atcgagcccc cacccgaatg a 1551

<210> 377
<211> 516
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(37)

<400> 377
Met Ser Gln Phe Leu Arg Phe Ala Val Arg Arg Ile Cys Thr Gln Trp
 1   5   10   15
Trp Cys Arg Thr Thr Gly Ile Ile Leu Leu Val Val Phe Ala Gly Val
            20   25   30
Ser Ala Ala Leu Gly Gln Ala Gly Phe Asp Asp Asp Arg Val Met Leu
        35   40   45
Gln Gly Phe Tyr Trp Glu Ser Tyr Arg His Gly His Leu Glu Lys Phe
    50   55   60
Pro Asn Tyr Gly Asp Lys His Trp Tyr Val Ile Val Gln Asn Leu Ala
65   70   75   80
Pro Glu Ile Arg Ala Gly Arg Phe Asn Leu Ile Trp Leu Pro Pro Pro
                85   90   95
Ser Tyr Ala Gly Gly Phe Ser Ala Gly Tyr Asn Pro Lys Glu Tyr Phe
            100   105   110
Arg Leu Asp Asn Ser Tyr Gly Thr Phe Asp Glu His Arg Ala Ala Leu
        115   120   125
Glu Ala Leu Leu Arg Asn Gly Ile Glu Pro Val Ala Asp Leu Val Ile
    130   135   140
Asn His Arg Asp Gly Asn Gln Ala Trp Ala Asp Phe Lys Asn Pro Asp
145   150   155   160
Trp Gly Thr Trp Ala Ile Cys Arg Thr Asp Glu Ala Phe Thr Asn Val
                165   170   175
Asn Ser Gly Ile Thr Asn Thr Pro Asn Asp Gln Arg Gly Asn Cys Glu
            180   185   190
Glu Glu Ala Asp Tyr Arg Pro Glu Gly Thr Tyr Asn Tyr Ser Asp Phe
        195   200   205
Arg Asp Ile Ala His Ala Asp Pro Arg Val Arg Arg Asp Ile Glu Arg
    210   215   220
Tyr Leu Leu Gln Leu Lys Ser Leu Gly Tyr Arg Gly Trp Arg Tyr Asp
225   230   235   240
Met Val His Gly Tyr Asn Ala Arg Trp Val Ala Leu Tyr Asn Arg Leu
                245   250   255
Thr Gln Pro Thr Phe Ser Val Gly Glu Tyr Asp Trp Asp Lys His Ala
            260   265   270
Gln Gln Arg Gly Trp Ile Trp Gly Thr Ala Thr Asn Pro Thr Pro Thr
        275   280   285
Gly Ala Asp His Leu Arg Thr Ser Ser Ser Val Phe Asp Phe Thr Ser
    290   295   300
Gln Phe Ser Leu Lys Ser Ile Asn Ser Gly Asn Tyr Thr Ala Leu Tyr
305   310   315   320
Gly Phe Gly Asn Gly Ile Gly Leu Val Gly Asp Thr Thr Asp Gly Leu
                325   330   335
Pro Trp Lys Asn Arg Ala Val Thr Phe Leu Glu Asn His Asp Thr Gly
            340   345   350
Tyr Arg Thr Asn Glu Asp Gly Thr Pro Gln Gln Gly His Gln Phe Asp
        355   360   365
Ser Phe Ala Asn Asn Trp Gln Val Glu Gln Gly Tyr Ala Gln Ile Leu
    370   375   380
Thr His Pro Gly Val Pro Thr Val Tyr Trp Lys His Tyr Phe Asp Trp
385   390   395   400
Gly Ser Asp Leu Gln Asn Lys Ile Arg Ala Leu Ile Asn Ala Arg Lys

405 410 415
Val Ala Gly Val Asn Ala Gly Ser Lys Val Asp Pro Gln Asp Asn Ala
420 425 430
Lys Tyr Ser Gly Val Tyr Ala Ala Arg Ile Ser Gly Arg Asn Gly Glu
435 440 445
Leu Tyr Val Arg Ile Gly Gly Ser Asp Gln Gln Trp Gln Pro Ser Ala
450 455 460
Ser Gly Tyr Ala Asp Tyr Arg Glu Tyr Ala Gln Gly Ala Gly Trp Lys
465 470 475 480
Val Trp Val Lys Leu Pro Gly Asn Pro Pro Val Gln Gln Val Pro Leu
485 490 495
Pro Ser Ala Leu Pro Val Pro Gln Tyr Arg Pro Pro Asp Gln Ile Glu
500 505 510
Pro Pro Pro Glu
515

<210> 378
<211> 1269
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 378
atgttgaagc agttcacgaa gcgcctgatc accctgacga gcctgctggc gctcgtcctc 60
gtcgcaccgt tggccagcgc gggcccgctg gatggcaaca gcagcgacgt catgttgcag 120
ggcttccatt ggtactcgta ccagtcgttt ccgtggtggg gcgtcatcaa gaacaacgcg 180
gcgagcatca aggccgacgg cttcaccatg gtgtggctgc cgccgcccag cgacgcggcc 240
tccaacgagg gctacctgcc gcgccggctc gagctgctgg acagcaagta tggcacccgg 300
acggacctgg tcaacgccct gtccgcgctg aatgccaatg gtgtgaagcc cattgcggac 360
atcgtcatca accaccgcgt gggcaccacg ggctgggcgg acttcacgct gcctccgtgg 420
ggctcgaacg cggtgtgccg cggcgacgag tggagcgggg ccacgggcaa cgcggatacg 480
ggcgatggct tcaacgccgg gcgcgacatc gatcacacgc agaccttcgt gcaggacggc 540
atcgtcacct ggatgaacaa ctcgctgaag agcgtcgggt tcgcgggttg gcggtatgac 600
tacgtgaagg gctacagcgg ctcctacgtc ggctcgtaca acacccgcac gacgccgtac 660
ttctccgtgg gcgagctgtg gacggacctg gacctgaaca accccaaccc ccaccgccag 720
ctgatcatga attggatcga cgcgacgggt ggccggtccg cggcgttcga cttcacgacc 780
aagggcctgc tgcagcaggc ggtgcagtac aacgagttct ggcggctgaa ggatgcggcg 840
ggcgcgccag cgggtgccat tggttggtgg gcagcgaagt ccgtgacctt catcgacaat 900
cacgacacgg gcccgagcta tccgagcggc ggccagaacc actggccgtt ccctggtgac 960
aagatcctcc aggggtacgc ctacatcctg actcactctg gcatcccctg cgtgtactgg 1020
gtgcactaca aggactgggg ccaggcgaac acggacgcca tcaagaagct gatcagcatc 1080
cgcaagtcca agggcatcac cagcacctcc tcggtgagca tccaggccgc ggacagctcg 1140
aagtacgccg ccatcatcac cggcaacaac ggcaaggtgg ccgtgaagat cggcttcggc 1200
gcctggtctc cgccgggcac ctggacgctg gccacctccg gcaacaacta cgccgtctgg 1260
acgcagtaa 1269

<210> 379
<211> 422
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(27)

<400> 379
Met Leu Lys Gln Phe Thr Lys Arg Leu Ile Thr Leu Thr Ser Leu Leu
1 5 10 15
Ala Leu Val Leu Val Ala Pro Leu Ala Ser Ala Gly Pro Leu Asp Gly
20 25 30
Asn Ser Ser Asp Val Met Leu Gln Gly Phe His Trp Tyr Ser Tyr Gln
35 40 45
Ser Phe Pro Trp Trp Gly Val Ile Lys Asn Asn Ala Ala Ser Ile Lys
50 55 60

Ala Asp Gly Phe Thr Met Val Trp Leu Pro Pro Pro Ser Asp Ala Ala
65 70 75 80
Ser Asn Glu Gly Tyr Leu Pro Arg Arg Leu Glu Leu Leu Asp Ser Lys
85 90 95
Tyr Gly Thr Arg Thr Asp Leu Val Asn Ala Leu Ser Ala Leu Asn Ala
100 105 110
Asn Gly Val Lys Pro Ile Ala Asp Ile Val Ile Asn His Arg Val Gly
115 120 125
Thr Thr Gly Trp Ala Asp Phe Thr Leu Pro Pro Trp Gly Ser Asn Ala
130 135 140
Val Cys Arg Gly Asp Glu Trp Ser Gly Ala Thr Gly Asn Ala Asp Thr
145 150 155 160
Gly Asp Gly Phe Asn Ala Gly Arg Asp Ile Asp His Thr Gln Thr Phe
165 170 175
Val Gln Asp Gly Ile Val Thr Trp Met Asn Asn Ser Leu Lys Ser Val
180 185 190
Gly Phe Ala Gly Trp Arg Tyr Asp Tyr Val Lys Gly Tyr Ser Gly Ser
195 200 205
Tyr Val Gly Ser Tyr Asn Thr Arg Thr Thr Pro Tyr Phe Ser Val Gly
210 215 220
Glu Leu Trp Thr Asp Leu Asp Leu Asn Asn Pro Asn Pro His Arg Gln
225 230 235 240
Leu Ile Met Asn Trp Ile Asp Ala Thr Gly Gly Arg Ser Ala Ala Phe
245 250 255
Asp Phe Thr Thr Lys Gly Leu Leu Gln Gln Ala Val Gln Tyr Asn Glu
260 265 270
Phe Trp Arg Leu Lys Asp Ala Ala Gly Ala Pro Ala Gly Ala Ile Gly
275 280 285
Trp Trp Ala Ala Lys Ser Val Thr Phe Ile Asp Asn His Asp Thr Gly
290 295 300
Pro Ser Tyr Pro Ser Gly Gly Gln Asn His Trp Pro Phe Pro Gly Asp
305 310 315 320
Lys Ile Leu Gln Gly Tyr Ala Tyr Ile Leu Thr His Ser Gly Ile Pro
325 330 335
Cys Val Tyr Trp Val His Tyr Lys Asp Trp Gly Gln Ala Asn Thr Asp
340 345 350
Ala Ile Lys Lys Leu Ile Ser Ile Arg Lys Ser Lys Gly Ile Thr Ser
355 360 365
Thr Ser Ser Val Ser Ile Gln Ala Ala Asp Ser Ser Lys Tyr Ala Ala
370 375 380
Ile Ile Thr Gly Asn Asn Gly Lys Val Ala Val Lys Ile Gly Phe Gly
385 390 395 400
Ala Trp Ser Pro Pro Gly Thr Trp Thr Leu Ala Thr Ser Gly Asn Asn
405 410 415
Tyr Ala Val Trp Thr Gln
420

<210> 380
<211> 1644
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 380
atgaagccgc gttgggcagc gtggcttgtc ctcgtatgct tgctggcggg ctgcttcggg 60
ggagaaagcg cgcctgccga tcatgcggcc cggtcttccg ctccggacgg aacgccggcg 120
ccggtcccgt ccgctgcagc cgcagggtgg aaagcggatg aacaaccgtc gacggtgttt 180
tacgagattt tcgtgcgctc gttctacgac tccgacggag acggcatcgg ggatttgaac 240
ggcgtgacgg agaagctcga ctatttgaag gagttgggag tcggcggcat ctggctgatg 300
ccgatcaacg cgtcgccgag ctaccatggc tacgatgtga cggactacta tgcgatcaac 360
cccgactacg gcacgttgga agacttgaaa aggctgctgg ccgaagcgca tgcgcggggc 420
atcaaagtca ttatggatct tgtcgtcaac catacgagcc gcgagcatcc ctggttcaag 480
gaggcgcttg ccgatccgga cagcccttac cgcaactggt atacgatcaa accgtcggac 540
gaggcggcgc cttcggacaa cgcggcgggg accggcagtc cctggcatga tgcgggcagc 600
tacaaatacc tcgggatttt ctgggaaggc atgcccgatt tgaatttcga cgagccgaag 660
gttcgcgaag agatgatcaa ggtcggccgg tactggctgg agcagggggt tgacggcttc 720

```
cggctggatg cggccaagca tatttacggg gatttcgcct ccaccgtcca ctcggaggaa     780
atcgcgggca aaaacaaagc gtggtggcag gaattccgca gcgggttgaa cgaggccaat     840
ccgaacgctt acctcgtggg cgaaatctgg gattcggcga cgacgatcgc gcctctgctc     900
gacaaggcgc tcgattccgg attcaacttc gatttggcgg gacggattct cgacgccgcg     960
aaggggggaac gcgacccgga cgtcgccttc agcctgaagc gggtatacgg cgcttacggc   1020
caagcgtccg gcggttcgtt cacggacgcg acgtttctct ccaatcacga ccagaaccgc   1080
gccatgagcg tgctcggcgg gaacgtaaat cacgcgaaga tggccgcggc ggtgctgctg   1140
acgctgccgg gcaatccttt cctttattac ggggaggaaa tcggcatgct cggcatgaag   1200
cccgatgaag cgatccgcga gcccatgcct tggtacgccg gcccatcggg cgggccgggg   1260
cagacgtctt gggaaccgct gcggtacaac aagcccggcg ccacatcggt cgagcagcag   1320
gagcgggacg agcggtcgct gctccatcat taccgcacgc tgatcaagtg gcggaatgaa   1380
atgcccgagc tgcgggacgg ggatatcgac gagtacgata ccggcaacgc gaagctcggc   1440
gcttacgtgc ggatgacgaa ggacgcccgc gtcctcgtcg tgcacaatat gagcggcgag   1500
aagcagtcgg cgaagctgac ggcgtcggag cggtacggag cgttccggga aatcgtgcgc   1560
cagaccggca gcggcggcga aggcgcttcg cttgacgggg aaacgctctc catcgctccg   1620
tacagcacgg tcatcatccg ctag                                          1644

<210> 381
<211> 547
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(29)

<400> 381
Met Lys Pro Arg Trp Ala Ala Trp Leu Val Leu Val Cys Leu Leu Ala
 1               5                  10                  15
Gly Cys Phe Gly Gly Glu Ser Ala Pro Ala Asp His Ala Ala Arg Ser
            20                  25                  30
Ser Ala Pro Asp Gly Thr Pro Ala Pro Val Pro Ser Ala Ala Ala Ala
        35                  40                  45
Gly Trp Lys Ala Asp Glu Gln Pro Ser Thr Val Phe Tyr Glu Ile Phe
    50                  55                  60
Val Arg Ser Phe Tyr Asp Ser Asp Gly Asp Gly Ile Gly Asp Leu Asn
65                  70                  75                  80
Gly Val Thr Glu Lys Leu Asp Tyr Leu Lys Glu Leu Gly Val Gly Gly
                85                  90                  95
Ile Trp Leu Met Pro Ile Asn Ala Ser Pro Ser Tyr His Gly Tyr Asp
            100                 105                 110
Val Thr Asp Tyr Tyr Ala Ile Asn Pro Asp Tyr Gly Thr Leu Glu Asp
        115                 120                 125
Leu Lys Arg Leu Leu Ala Glu Ala His Ala Arg Gly Ile Lys Val Ile
    130                 135                 140
Met Asp Leu Val Val Asn His Thr Ser Arg Glu His Pro Trp Phe Lys
145                 150                 155                 160
Glu Ala Leu Ala Asp Pro Asp Ser Pro Tyr Arg Asn Trp Tyr Thr Ile
                165                 170                 175
Lys Pro Ser Asp Glu Ala Ala Pro Ser Asp Asn Ala Ala Gly Thr Gly
            180                 185                 190
Ser Pro Trp His Asp Ala Gly Ser Tyr Lys Tyr Leu Gly Ile Phe Trp
        195                 200                 205
Glu Gly Met Pro Asp Leu Asn Phe Asp Glu Pro Lys Val Arg Glu Glu
    210                 215                 220
Met Ile Lys Val Gly Arg Tyr Trp Leu Glu Gln Gly Val Asp Gly Phe
225                 230                 235                 240
Arg Leu Asp Ala Ala Lys His Ile Tyr Gly Asp Phe Ala Ser Thr Val
                245                 250                 255
His Ser Glu Glu Ile Ala Gly Lys Asn Lys Ala Trp Trp Gln Glu Phe
            260                 265                 270
Arg Ser Gly Leu Asn Glu Ala Asn Pro Asn Ala Tyr Leu Val Gly Glu
        275                 280                 285
Ile Trp Asp Ser Ala Thr Thr Ile Ala Pro Leu Leu Asp Lys Ala Leu
    290                 295                 300
Asp Ser Gly Phe Asn Phe Asp Leu Ala Gly Arg Ile Leu Asp Ala Ala
```

305 310 315 320
Lys Gly Glu Arg Asp Pro Asp Val Ala Phe Ser Leu Lys Arg Val Tyr
325 330 335
Gly Ala Tyr Gly Gln Ala Ser Gly Gly Ser Phe Thr Asp Ala Thr Phe
340 345 350
Leu Ser Asn His Asp Gln Asn Arg Ala Met Ser Val Leu Gly Gly Asn
355 360 365
Val Asn His Ala Lys Met Ala Ala Ala Val Leu Leu Thr Leu Pro Gly
370 375 380
Asn Pro Phe Leu Tyr Tyr Gly Glu Glu Ile Gly Met Leu Gly Met Lys
385 390 395 400
Pro Asp Glu Ala Ile Arg Glu Pro Met Pro Trp Tyr Ala Gly Pro Ser
405 410 415
Gly Gly Pro Gly Gln Thr Ser Trp Glu Pro Leu Arg Tyr Asn Lys Pro
420 425 430
Gly Ala Thr Ser Val Glu Gln Gln Glu Arg Asp Glu Arg Ser Leu Leu
435 440 445
His His Tyr Arg Thr Leu Ile Lys Trp Arg Asn Glu Met Pro Glu Leu
450 455 460
Arg Asp Gly Asp Ile Asp Glu Tyr Asp Thr Gly Asn Ala Lys Leu Gly
465 470 475 480
Ala Tyr Val Arg Met Thr Lys Asp Ala Arg Val Leu Val Val His Asn
485 490 495
Met Ser Gly Glu Lys Gln Ser Ala Lys Leu Thr Ala Ser Glu Arg Tyr
500 505 510
Gly Ala Phe Arg Glu Ile Val Arg Gln Thr Gly Ser Gly Gly Glu Gly
515 520 525
Ala Ser Leu Asp Gly Glu Thr Leu Ser Ile Ala Pro Tyr Ser Thr Val
530 535 540
Ile Ile Arg
545

<210> 382
<211> 4176
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 382
gtgctgctgc gcacgcccac cgcgcgcaga ggactggccg tcgccacggc ggccggtgtc 60
ctcgcgacgc tgatccacgc ggcggcggcc ccggccgccc aggcggcccc cccgccgccg 120
accgacgccg aactggccgc cgagccgacc cggcacgccc tcaccaccga gcagttctac 180
ttcgtgctcc cggaccgctt cgccaacggc gacccgtcca acgacaccgc cggcatcccc 240
ggcggcccgc tcgaccacgg gtacgacccc gacgacgagg gcttctacca gggcggcgac 300
ctgcgcggcc tgatcgagaa cctggactac atcgagggcc tgggcaccac cgccctctgg 360
atggcgcccg tcttcgagaa caagccggtc cagggcgagg gggaccacgc ctcggccggc 420
taccacggct actggatcac cgacttcacc cgggtggacc cgcacttcgg caccaacgag 480
gaactcgccg aactggtcgg ggccgcgcac gagcgggaca tcaaggtctt cttcgacgtc 540
atcaccaacc acaccgccga cgtcatccac tacgacgagg gcgcctacga ctacctctcc 600
aagggcgcct tcccgtacct cgacaccgag ggccgccctt tcgacgacct cgaccacgcc 660
ggtgccgagg acgggggttt ccccgaactc gacaccgcct ccttcccgta cacgccggtg 720
acctccgacg ccgaggccga cctcaaggtc cccgactggc tcaacgaccc gacgctctac 780
cacaaccgcg gcgactcctc cttcgccggc gagagttccc tgtacggcga cttccacggc 840
ctggacgacc tgttcaccgc ccggcccgag gtcgtcgagg gcatggtgga gatctaccgc 900
acctgggtcg aggacttcgg catcgacggt ttccgcatcg acaccgtcaa gcacgtcaac 960
atggagttct ggacccgctg gtccgccgcc atgaaggagc acgccgcgtc cctcggccac 1020
gaggacttct tcatgttcgg cgaggtgtac tcctccgatg tcgccgagaa ggcccggtac 1080
gtcaccgagg gcggcctgaa cgccaccctg gacttcgcct tccaggaggc cgcccgctcc 1140
ttcgcctcgc agggcggctc cgcgaagcgc ctgtccgccc tcttcggcga ggaccaccgc 1200
tacaccaccg ccgacaccaa cgcccagcag tcggtcacct tcctgggcaa ccacgacatg 1260
ggccggatcg gtcacttcct gcgcgccgac aacccggacg ccgacgaggc cgagctgctc 1320
gagcgttccc gcctcgccca cgagctgatg ttcctgtccc ggggcaaccc ggtgatctac 1380
tacggcgacg agcagggctt caccggttcc ggtggcgaca aggccgcccg tgccaccctc 1440
ttcggctcca cgacgccgct ctacctggag gacccgcaac tgggcaccga ccgcacccac 1500
gcggaggacg cccacgacac cgaccatccc ctgtaccggg cgatcgccgc cctctccgcc 1560
ctcacccgcg agcacccggc gctgcgcgac ggcgcgcaga ccgagcggta cgcggaggac 1620

```
ggcgccggcg tccgggccca cacccgcacc gacgccgacg aacgggtgga gtacctggtc    1680
gccgccaaca acgcggagga cggacgcacc gtcgccgtgc ccaccgacac ccccggcgcc    1740
gtcttcgagg tcctgtacgg cggggtcggc gacggttccg cgaccccggt gaccgccgac    1800
gccgacggcg tcgtcgagct gaccgtgccc gccctgtcca ccgcggtact gcgcgccgac    1860
cgacccctgc ccgacccgca ggccgccccc gccgtggagc tcaccgcccc cgacgccggc    1920
gcacacggtg tcgtggagct ctccgccgcc gtcaccggcg gcgcgctcaa ccgcgtcgcc    1980
ttcgccgcgc aggtcggcga cggcgactgg gaggtgctcg gggtctccga cgccccgccg    2040
caccgcgtca cccacaccgt cccggccggc acctccgccg gtaccccgct gcgctaccgc    2100
gccgtcgccg tggattcctc cggcgccacc gccgcgaccg gcgcggagac caccgccggc    2160
gccccgccgc cggaacgggc gccgcgcgcc gtgagccggg agtgggccgt cgtccactac    2220
caccgccccg acgggaacta cgacgactgg tcgctgtggg cgtggggcga cctccacgag    2280
agcgagagct tcgaggactg gccccacggc cggcccttcg tcggccgcga cgccttcggc    2340
gccttcgccc gggtccgcct ggcccccgac gcctccgaac tcggctacct gatcgtggac    2400
ggcgacggca ccaaggacgg tgacctcgac cgggtcatcg acctctccga gaccggcgag    2460
gtctggatcg agcagggcga gccggaggcg ctcaccgagg cgcccgaagg cgtgtatccg    2520
cccgccgacg agtcgcgcgc cgtcctgcgg taccaccgac ccgacggcga ctacgatgga    2580
tggggcgtgc acgtgtggac cggcgcggcc gagcccaccg actggagccg acccgttccg    2640
gcctccgagg tggactcctt cggcgccgtc tacgagatcc cgctcgccga gggcgcgacc    2700
tctctgagct acatcatcca ccggggcgac gagaaggacc tgcccgccga ccaatccctg    2760
gacctgcggg ccgtcggtta cgaggtgtgg atcatctccg gagaggagac gtacctgttg    2820
cccagcggcg gagccgcccc cgacctgaac cccgaccgcg cccgggccca ctggatcgac    2880
gccgagaccg tcgtgtggcc cggcgttgcc ccgcccggtg gcagccggca gctgatccac    2940
caccccgagg gcgccatcgc cctcgaggac ggcgcgctca ccgacgaggg ccactggatc    3000
accactccgc cggcgggcga gcgcgacccg cgaccgggtt tgaccggcct tgagccatcg    3060
gacatgggtc gctggtttgc cgagcgcggc cagcccggct accgcgcccg ccagctggcc    3120
gatcacgtct ggcaggcgtg cgccatccgg gcggacgagc tgcggaccct gccggccgcg    3180
ttgcgcacag acatagaaga ctcgttccgc gtggacacgt tggtcgagac ggacgttcgg    3240
ccggccgacg cgggcctgac cgagaaggct ctccatcggc tcgacgacgg ccggctcatc    3300
gagtcggtat tgatgcgcta tcccgcgcgc gggccgcggc gctggcgggc aacggcgtgc    3360
atctcaagcc aggctggctg cgcagtaggc tgcccattct gcgcgactgg cgagcttggc    3420
ttcatgcgcg acctcgaggt gtctgagatc gtcgaccagg cgcgctactg gcgccgccgc    3480
ctggccgccg agggacgccg gctgaccaat ctcgtgttca tgggcatggg cgaaccgctg    3540
ctcaaccccg accgcgtgct ggccgctgcg gaagcgctgt ccgacgcgcg ccgtttcggg    3600
cttggcgcgc gccacatcac gctcagcacc agcggcgtcg tgccggccat ggagcggctg    3660
attgccctcc ggccccagta caccctcgcc gtttcgctgc acgctgcccg gcccgaccta    3720
cgcgacgtgc tggtgccgct caaccgccgc tacccggttg gcgacgtggt agccgccgcc    3780
agcgcgtatg cgcgcctgac cggccgccgg gtgagctacg agtacgtgat gatcgactcg    3840
gtcaacgaca cgccgcgcga cgcagccgag ctgacccagc tgctcggtgg ccgccatgcc    3900
cacgtcaact tgattccgat gaaccccgtc gcgcacacgc catggcaggc cagccccact    3960
tcccggatcg ccgagttcgc gtcgcttctg cgctcgggcg gcatccgcac gaccgtgcgc    4020
cgcaatcgcg gcttggagat aggtgctgca tgcgggcagc tggctgccga cagcgcaggc    4080
gcgccgccac cggccgcggt ggcgcgacgt cgcgagctgc tggtgctcaa cagcgcagcg    4140
gcgttggccg cacgaccggc tgtttcgagg agctga                              4176

<210> 383
<211> 1391
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(35)

<400> 383
Val Leu Leu Arg Thr Pro Thr Ala Arg Arg Gly Leu Ala Val Ala Thr
 1               5                  10                  15
Ala Ala Gly Val Leu Ala Thr Leu Ile His Ala Ala Ala Ala Pro Ala
            20                  25                  30
Ala Gln Ala Ala Pro Pro Pro Pro Thr Asp Ala Glu Leu Ala Ala Glu
        35                  40                  45
Pro Thr Arg His Ala Leu Thr Thr Glu Gln Phe Tyr Phe Val Leu Pro
    50                  55                  60
Asp Arg Phe Ala Asn Gly Asp Pro Ser Asn Asp Thr Ala Gly Ile Pro
65                  70                  75                  80
Gly Gly Pro Leu Asp His Gly Tyr Asp Pro Asp Asp Glu Gly Phe Tyr
                85                  90                  95
```

```
Gln Gly Gly Asp Leu Arg Gly Leu Ile Glu Asn Leu Asp Tyr Ile Glu
            100                 105                 110
Gly Leu Gly Thr Thr Ala Leu Trp Met Ala Pro Val Phe Glu Asn Lys
        115                 120                 125
Pro Val Gln Gly Glu Gly Asp His Ala Ser Ala Gly Tyr His Gly Tyr
    130                 135                 140
Trp Ile Thr Asp Phe Thr Arg Val Asp Pro His Phe Gly Thr Asn Glu
145                 150                 155                 160
Glu Leu Ala Glu Leu Val Gly Ala Ala His Glu Arg Asp Ile Lys Val
                165                 170                 175
Phe Phe Asp Val Ile Thr Asn His Thr Ala Asp Val Ile His Tyr Asp
            180                 185                 190
Glu Gly Ala Tyr Asp Tyr Leu Ser Lys Gly Ala Phe Pro Tyr Leu Asp
        195                 200                 205
Thr Glu Gly Arg Pro Phe Asp Asp Leu Asp His Ala Gly Ala Glu Asp
    210                 215                 220
Gly Gly Phe Pro Glu Leu Asp Thr Ala Ser Phe Pro Tyr Thr Pro Val
225                 230                 235                 240
Thr Ser Asp Ala Glu Ala Asp Leu Lys Val Pro Asp Trp Leu Asn Asp
                245                 250                 255
Pro Thr Leu Tyr His Asn Arg Gly Asp Ser Ser Phe Ala Gly Glu Ser
            260                 265                 270
Ser Leu Tyr Gly Asp Phe His Gly Leu Asp Asp Leu Phe Thr Ala Arg
        275                 280                 285
Pro Glu Val Val Glu Gly Met Val Glu Ile Tyr Arg Thr Trp Val Glu
    290                 295                 300
Asp Phe Gly Ile Asp Gly Phe Arg Ile Asp Thr Val Lys His Val Asn
305                 310                 315                 320
Met Glu Phe Trp Thr Arg Trp Ser Ala Ala Met Lys Glu His Ala Ala
                325                 330                 335
Ser Leu Gly His Glu Asp Phe Phe Met Phe Gly Glu Val Tyr Ser Ser
            340                 345                 350
Asp Val Ala Glu Lys Ala Arg Tyr Val Thr Glu Gly Gly Leu Asn Ala
        355                 360                 365
Thr Leu Asp Phe Ala Phe Gln Glu Ala Ala Arg Ser Phe Ala Ser Gln
    370                 375                 380
Gly Gly Ser Ala Lys Arg Leu Ser Ala Leu Phe Gly Glu Asp His Arg
385                 390                 395                 400
Tyr Thr Thr Ala Asp Thr Asn Ala Gln Gln Ser Val Thr Phe Leu Gly
                405                 410                 415
Asn His Asp Met Gly Arg Ile Gly His Phe Leu Arg Ala Asp Asn Pro
            420                 425                 430
Asp Ala Asp Glu Ala Glu Leu Leu Glu Arg Ser Arg Leu Ala His Glu
        435                 440                 445
Leu Met Phe Leu Ser Arg Gly Asn Pro Val Ile Tyr Tyr Gly Asp Glu
    450                 455                 460
Gln Gly Phe Thr Gly Ser Gly Gly Asp Lys Ala Ala Arg Ala Thr Leu
465                 470                 475                 480
Phe Gly Ser Thr Thr Pro Leu Tyr Leu Glu Asp Pro Gln Leu Gly Thr
                485                 490                 495
Asp Arg Thr His Ala Glu Asp Ala His Asp Thr Asp His Pro Leu Tyr
            500                 505                 510
Arg Ala Ile Ala Ala Leu Ser Ala Leu Thr Arg Glu His Pro Ala Leu
        515                 520                 525
Arg Asp Gly Ala Gln Thr Glu Arg Tyr Ala Glu Asp Gly Ala Gly Val
    530                 535                 540
Arg Ala His Thr Arg Thr Asp Ala Asp Glu Arg Val Glu Tyr Leu Val
545                 550                 555                 560
Ala Ala Asn Asn Ala Glu Asp Gly Arg Thr Val Ala Val Pro Thr Asp
                565                 570                 575
Thr Pro Gly Ala Val Phe Glu Val Leu Tyr Gly Gly Val Gly Asp Gly
            580                 585                 590
Ser Ala Thr Pro Val Thr Ala Asp Ala Asp Gly Val Val Glu Leu Thr
        595                 600                 605
Val Pro Ala Leu Ser Thr Ala Val Leu Arg Ala Asp Arg Pro Leu Pro
    610                 615                 620
Asp Pro Gln Ala Ala Pro Ala Val Glu Leu Thr Ala Pro Asp Ala Gly
625                 630                 635                 640
```

```
Ala His Gly Val Val Glu Leu Ser Ala Ala Val Thr Gly Gly Ala Leu
                645                 650                 655
Asn Arg Val Ala Phe Ala Ala Gln Val Gly Asp Gly Asp Trp Glu Val
            660                 665                 670
Leu Gly Val Ser Asp Ala Pro Pro His Arg Val Thr His Thr Val Pro
        675                 680                 685
Ala Gly Thr Ser Ala Gly Thr Pro Leu Arg Tyr Arg Ala Val Ala Val
    690                 695                 700
Asp Ser Ser Gly Ala Thr Ala Ala Thr Gly Ala Glu Thr Thr Ala Gly
705                 710                 715                 720
Ala Pro Pro Pro Glu Arg Ala Pro Arg Ala Val Ser Arg Glu Trp Ala
                725                 730                 735
Val Val His Tyr His Arg Pro Asp Gly Asn Tyr Asp Asp Trp Ser Leu
            740                 745                 750
Trp Ala Trp Gly Asp Leu His Glu Ser Glu Ser Phe Glu Asp Trp Pro
        755                 760                 765
His Gly Arg Pro Phe Val Gly Arg Asp Ala Phe Gly Ala Phe Ala Arg
    770                 775                 780
Val Arg Leu Ala Pro Asp Ala Ser Glu Leu Gly Tyr Leu Ile Val Asp
785                 790                 795                 800
Gly Asp Gly Thr Lys Asp Gly Asp Leu Asp Arg Val Ile Asp Leu Ser
                805                 810                 815
Glu Thr Gly Glu Val Trp Ile Glu Gln Gly Glu Pro Glu Ala Leu Thr
            820                 825                 830
Glu Ala Pro Glu Gly Val Tyr Pro Pro Ala Asp Glu Ser Arg Ala Val
        835                 840                 845
Leu Arg Tyr His Arg Pro Asp Gly Asp Tyr Asp Gly Trp Gly Val His
    850                 855                 860
Val Trp Thr Gly Ala Ala Glu Pro Thr Asp Trp Ser Arg Pro Val Pro
865                 870                 875                 880
Ala Ser Glu Val Asp Ser Phe Gly Ala Val Tyr Glu Ile Pro Leu Ala
                885                 890                 895
Glu Gly Ala Thr Ser Leu Ser Tyr Ile Ile His Arg Gly Asp Glu Lys
            900                 905                 910
Asp Leu Pro Ala Asp Gln Ser Leu Asp Leu Arg Ala Val Gly Tyr Glu
        915                 920                 925
Val Trp Ile Ile Ser Gly Glu Glu Thr Tyr Leu Leu Pro Ser Gly Gly
    930                 935                 940
Ala Ala Pro Asp Leu Asn Pro Asp Arg Ala Arg Ala His Trp Ile Asp
945                 950                 955                 960
Ala Glu Thr Val Val Trp Pro Gly Val Ala Pro Pro Gly Gly Ser Arg
                965                 970                 975
Gln Leu Ile His His Pro Glu Gly Ala Ile Ala Leu Glu Asp Gly Ala
            980                 985                 990
Leu Thr Asp Glu Gly His Trp Ile Thr Thr Pro Pro Ala Gly Glu Arg
        995                 1000                1005
Asp Pro Arg Pro Gly Leu Thr Gly Leu Glu Pro Ser Asp Met Gly Arg
    1010                1015                1020
Trp Phe Ala Glu Arg Gly Gln Pro Gly Tyr Arg Ala Arg Gln Leu Ala
1025                1030                1035                1040
Asp His Val Trp Gln Ala Cys Ala Ile Arg Ala Asp Glu Leu Arg Thr
                1045                1050                1055
Leu Pro Ala Ala Leu Arg Thr Asp Ile Glu Asp Ser Phe Arg Val Asp
            1060                1065                1070
Thr Leu Val Glu Thr Asp Val Arg Pro Ala Asp Ala Gly Leu Thr Glu
        1075                1080                1085
Lys Ala Leu His Arg Leu Asp Asp Gly Arg Leu Ile Glu Ser Val Leu
    1090                1095                1100
Met Arg Tyr Pro Ala Arg Gly Pro Arg Arg Trp Arg Ala Thr Ala Cys
1105                1110                1115                1120
Ile Ser Ser Gln Ala Gly Cys Ala Val Gly Cys Pro Phe Cys Ala Thr
                1125                1130                1135
Gly Glu Leu Gly Phe Met Arg Asp Leu Glu Val Ser Glu Ile Val Asp
            1140                1145                1150
Gln Ala Arg Tyr Trp Arg Arg Arg Leu Ala Ala Glu Gly Arg Arg Leu
        1155                1160                1165
Thr Asn Leu Val Phe Met Gly Met Gly Glu Pro Leu Leu Asn Pro Asp
    1170                1175                1180
```

```
Arg Val Leu Ala Ala Ala Glu Ala Leu Ser Asp Ala Arg Arg Phe Gly
1185                1190                1195                1200
Leu Gly Ala Arg His Ile Thr Leu Ser Thr Ser Gly Val Val Pro Ala
                1205                1210                1215
Met Glu Arg Leu Ile Ala Leu Arg Pro Gln Tyr Thr Leu Ala Val Ser
            1220                1225                1230
Leu His Ala Ala Arg Pro Asp Leu Arg Asp Val Leu Val Pro Leu Asn
        1235                1240                1245
Arg Arg Tyr Pro Val Gly Asp Val Val Ala Ala Ala Ser Ala Tyr Ala
    1250                1255                1260
Arg Leu Thr Gly Arg Arg Val Ser Tyr Glu Tyr Val Met Ile Asp Ser
1265                1270                1275                1280
Val Asn Asp Thr Pro Arg Asp Ala Ala Glu Leu Thr Gln Leu Leu Gly
                1285                1290                1295
Gly Arg His Ala His Val Asn Leu Ile Pro Met Asn Pro Val Ala His
            1300                1305                1310
Thr Pro Trp Gln Ala Ser Pro Thr Ser Arg Ile Ala Glu Phe Ala Ser
        1315                1320                1325
Leu Leu Arg Ser Gly Gly Ile Arg Thr Thr Val Arg Arg Asn Arg Gly
    1330                1335                1340
Leu Glu Ile Gly Ala Ala Cys Gly Gln Leu Ala Ala Asp Ser Ala Gly
1345                1350                1355                1360
Ala Pro Pro Pro Ala Ala Val Ala Arg Arg Arg Glu Leu Leu Val Leu
                1365                1370                1375
Asn Ser Ala Ala Ala Leu Ala Ala Arg Pro Ala Val Ser Arg Ser
            1380                1385                1390
```

<210> 384
<211> 1434
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 384

```
atgcacgcag gcgcaaacga ccgtacgata cgggttcccg tctctcgccc acctcggcgc      60
acgtggtgga tcgtcgcgct gctgctgctg gtgtcggcat cggcgcaggc cgatgcggta     120
ttgcacgcat tcaactggcg ctatgccacc gtcgaggcac gtgcaaacga gattgccgcg     180
ctcggttatc gcgccgtgct cgtcgccccg ccactgcgct cgcagggcaa cgaatggtgg     240
gcccgctacc agccgcagga ctaccgcgtc atcgagcacc cgctcggcga caccgacgcg     300
ttcgtgcgca tggagtccgc gctgcgccag gccggcgtgc ggctgtatgc cgacatcgtg     360
ctgaaccaca tggcgaacga atcgggctgg cgttctgatc tgaactaccc cggcacgctg     420
gtgctgaacc agtacgccgc caacccgagc cgctatgcgc gtctgcgcct gttcggcgac     480
ctgcgctaca acttcctgtc gcaacacgac ttcggcccgg cgaactgcat caccgactac     540
ggcagcgtgt ggcaggtgca gacctggcgg ctgtgcggcg ccgcgcccga tgtcggcctg     600
ccggatctgg tcgacaacga ctgggtggtg tcgcagcagc aacagtacct gcgctcgctg     660
aaggcgctcg gtgtcagcgg tttccgcatc gacgccgcca agcacatgcc gaacacgcac     720
atgaaccggg tgctgacgtc gaacatcaag cagggcgtgc acgtgttcgg cgagatcatc     780
accgacggcg gcgtcggcga cacgtcatac gaccgcttcc tggcgccgta tctcgcccag     840
accgaccacg gcgcgtacga cttcccgttg ttccaccaga tccggaacgc gttccggccc     900
ggcggcagca tgaacctgct ggtcgacccc ggtgcctacg ggcaggccct gccgcggcac     960
cgtgcgatca ccttcacggt cacgcatgac atccccaaca acggcatctt ccgctacctg    1020
atcctcgacc gcaccgacga gacgctggcc tatgcctatg tgctcggccg cgacggtggc    1080
gtgccgctga tctattccga caacaacgaa agcggcgaca accgctgggt gaacgcctat    1140
cggcgctcgg atctggcgca gatgatccgc tttcacaaca gcgtgcaggg cagcgacatg    1200
caggtgctgt cgtatggcag ctgccatctg ctgttccgcc gcggcaaccg cggcatcgtc    1260
ggcatcaaca agtgcggcag cagcgtgaat gcgaccgtga acatgcacgg cagcgtgctg    1320
tggtggtaca ccgactaccg cgacgtgctg ggctcgggcc atgtggtcaa catcggcagc    1380
ggcagctaca cgttcactct gcccgcgcgt gcggcgcgga tgtggttgcg ctga          1434
```

<210> 385
<211> 477
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(37)

<400> 385
Met His Ala Gly Ala Asn Asp Arg Thr Ile Arg Val Pro Val Ser Arg
 1   5   10   15
Pro Pro Arg Arg Thr Trp Trp Ile Val Ala Leu Leu Leu Leu Val Ser
            20   25   30
Ala Ser Ala Gln Ala Asp Ala Val Leu His Ala Phe Asn Trp Arg Tyr
        35   40   45
Ala Thr Val Glu Ala Arg Ala Asn Glu Ile Ala Ala Leu Gly Tyr Arg
    50   55   60
Ala Val Leu Val Ala Pro Pro Leu Arg Ser Gln Gly Asn Glu Trp Trp
65   70   75   80
Ala Arg Tyr Gln Pro Gln Asp Tyr Arg Val Ile Glu His Pro Leu Gly
                85   90   95
Asp Thr Asp Ala Phe Val Arg Met Glu Ser Ala Leu Arg Gln Ala Gly
            100   105   110
Val Arg Leu Tyr Ala Asp Ile Val Leu Asn His Met Ala Asn Glu Ser
        115   120   125
Gly Trp Arg Ser Asp Leu Asn Tyr Pro Gly Thr Leu Val Leu Asn Gln
    130   135   140
Tyr Ala Ala Asn Pro Ser Arg Tyr Ala Arg Leu Arg Leu Phe Gly Asp
145   150   155   160
Leu Arg Tyr Asn Phe Leu Ser Gln His Asp Phe Gly Pro Ala Asn Cys
                165   170   175
Ile Thr Asp Tyr Gly Ser Val Trp Gln Val Gln Thr Trp Arg Leu Cys
            180   185   190
Gly Ala Ala Pro Asp Val Gly Leu Pro Asp Leu Val Asp Asn Asp Trp
        195   200   205
Val Val Ser Gln Gln Gln Gln Tyr Leu Arg Ser Leu Lys Ala Leu Gly
    210   215   220
Val Ser Gly Phe Arg Ile Asp Ala Ala Lys His Met Pro Asn Thr His
225   230   235   240
Met Asn Arg Val Leu Thr Ser Asn Ile Lys Gln Gly Val His Val Phe
                245   250   255
Gly Glu Ile Ile Thr Asp Gly Gly Val Gly Asp Thr Ser Tyr Asp Arg
            260   265   270
Phe Leu Ala Pro Tyr Leu Ala Gln Thr Asp His Gly Ala Tyr Asp Phe
        275   280   285
Pro Leu Phe His Gln Ile Arg Asn Ala Phe Arg Pro Gly Gly Ser Met
    290   295   300
Asn Leu Leu Val Asp Pro Gly Ala Tyr Gly Gln Ala Leu Pro Arg His
305   310   315   320
Arg Ala Ile Thr Phe Thr Val Thr His Asp Ile Pro Asn Asn Gly Ile
                325   330   335
Phe Arg Tyr Leu Ile Leu Asp Arg Thr Asp Glu Thr Leu Ala Tyr Ala
            340   345   350
Tyr Val Leu Gly Arg Asp Gly Gly Val Pro Leu Ile Tyr Ser Asp Asn
        355   360   365
Asn Glu Ser Gly Asp Asn Arg Trp Val Asn Ala Tyr Arg Arg Ser Asp
    370   375   380
Leu Ala Gln Met Ile Arg Phe His Asn Ser Val Gln Gly Ser Asp Met
385   390   395   400
Gln Val Leu Ser Tyr Gly Ser Cys His Leu Leu Phe Arg Arg Gly Asn
                405   410   415
Arg Gly Ile Val Gly Ile Asn Lys Cys Gly Ser Ser Val Asn Ala Thr
            420   425   430
Val Asn Met His Gly Ser Val Leu Trp Trp Tyr Thr Asp Tyr Arg Asp
        435   440   445
Val Leu Gly Ser Gly His Val Val Asn Ile Gly Ser Gly Ser Tyr Thr
    450   455   460
Phe Thr Leu Pro Ala Arg Ala Ala Arg Met Trp Leu Arg
465   470   475

<210> 386
<211> 1458

<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 386
atgaaaaata gcaacgcctt aaaaaatatc ttattatcgt cagttatttt aacttcagga 60
ctactcgcgc aagccaatcc gcgcacagtg atggtgcatt tatttgaatg gaagtggaat 120
gatatcgctc gtgagtgcga aacgtactta ggcccaaacg gctttgctgc aattcaagtt 180
tcgccgccga acgagcatgc ggttgttacc gggcgcccgt ggtacgaacg ttaccaaccg 240
gtgtcgtata ttttaaattc acgcagcggc aaccgcgaag aatttattaa catggttaag 300
gcttgtaaaa aagcaggtgt ggatatttat gtcgatgcag tgattaacca tatgaccgga 360
actttagcgc acggtgaaga gcgtgtgggt atttcaggct cacgctttgg ccgttttgct 420
tatccggatt attcttacga tgactttcac cattgcggtc gcaacggtaa tgatcagatc 480
agtaattatg gtgacagata tgaagtgcaa aactgtgcgc tggttggtct tgcagattta 540
aacatcgaaa aagattatgt gcgcgaccgt atcgccagat atttaaatga attagtcgat 600
attggtgtgg ctggatttcg catagatgcc gccaagcata tgtcgactaa ctcgttaaat 660
gaaattatta aacgtgtaaa aggcaacgct tatatttatc aggaagtgat tgatcaaggc 720
ggtgagccga ttcaggccaa agagtatttt aataatggtg atgtgacaga gtttaaatat 780
tcggtggata tgtcgcgtgt gatgttaaac ggaaaactaa cctggttaaa tgataaggcc 840
cagtttggtg aaggctggaa ctacatgcca tcagataaag cggtggtgtt tattgataac 900
catgataacc aacgcggtca tggcggcggc ggacatattt taaactataa agacggacgt 960
aattatgagt tggccagtgt actgatgctt gcatggcctt acggttaccc acaggtgatg 1020
tcgagctttg cgtttaccga tccaaactcg tcaccgccgg ctgatagtga aggtaatacg 1080
aaagatacga attgcctttc tcgtgcagag gggccgggcg ctcaggatgg ttgggtgtgt 1140
gaacaccgtt ggccgatgat tcgtgaaatg gttaaattca gaaatgtgac cagcagtaag 1200
ttttatttat cgaactggtg gagtaacggt aataaccaaa ttgctttttc gcgcggtgat 1260
ttaggatttg ttgtaatcaa ccacgaagag caggtgttgc gtcaacgatt gcaaaccggt 1320
ttaactgcag gtgtctactg tgatgtttta agtcaaaatt gtgaacgcac aattactgtt 1380
aatgacaatg gtgaagctga aatcatcgtt acgggtcaaa atgcagcggc catccatgcc 1440
cagagtaaat tgaagtag 1458

<210> 387
<211> 485
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(25)

<400> 387
Met Lys Asn Ser Asn Ala Leu Lys Asn Ile Leu Leu Ser Ser Val Ile
 1   5   10   15
Leu Thr Ser Gly Leu Leu Ala Gln Ala Asn Pro Arg Thr Val Met Val
            20   25   30
His Leu Phe Glu Trp Lys Trp Asn Asp Ile Ala Arg Glu Cys Glu Thr
        35   40   45
Tyr Leu Gly Pro Asn Gly Phe Ala Ala Ile Gln Val Ser Pro Pro Asn
    50   55   60
Glu His Ala Val Val Thr Gly Arg Pro Trp Tyr Glu Arg Tyr Gln Pro
65   70   75   80
Val Ser Tyr Ile Leu Asn Ser Arg Ser Gly Asn Arg Glu Glu Phe Ile
                85   90   95
Asn Met Val Lys Ala Cys Lys Lys Ala Gly Val Asp Ile Tyr Val Asp
            100   105   110
Ala Val Ile Asn His Met Thr Gly Thr Leu Ala His Gly Glu Glu Arg
        115   120   125
Val Gly Ile Ser Gly Ser Arg Phe Gly Arg Phe Ala Tyr Pro Asp Tyr
    130   135   140
Ser Tyr Asp Asp Phe His His Cys Gly Arg Asn Gly Asn Asp Gln Ile
145   150   155   160
Ser Asn Tyr Gly Asp Arg Tyr Glu Val Gln Asn Cys Ala Leu Val Gly
                165   170   175
Leu Ala Asp Leu Asn Ile Glu Lys Asp Tyr Val Arg Asp Arg Ile Ala

```
            180                 185                 190
Arg Tyr Leu Asn Glu Leu Val Asp Ile Gly Val Ala Gly Phe Arg Ile
        195                 200                 205
Asp Ala Ala Lys His Met Ser Thr Asn Ser Leu Asn Glu Ile Ile Lys
    210                 215                 220
Arg Val Lys Gly Asn Ala Tyr Ile Tyr Gln Glu Val Ile Asp Gln Gly
225                 230                 235                 240
Gly Glu Pro Ile Gln Ala Lys Glu Tyr Phe Asn Asn Gly Asp Val Thr
                245                 250                 255
Glu Phe Lys Tyr Ser Val Asp Met Ser Arg Val Met Leu Asn Gly Lys
            260                 265                 270
Leu Thr Trp Leu Asn Asp Lys Ala Gln Phe Gly Glu Gly Trp Asn Tyr
        275                 280                 285
Met Pro Ser Asp Lys Ala Val Val Phe Ile Asp Asn His Asp Asn Gln
    290                 295                 300
Arg Gly His Gly Gly Gly Gly His Ile Leu Asn Tyr Lys Asp Gly Arg
305                 310                 315                 320
Asn Tyr Glu Leu Ala Ser Val Leu Met Leu Ala Trp Pro Tyr Gly Tyr
                325                 330                 335
Pro Gln Val Met Ser Ser Phe Ala Phe Thr Asp Pro Asn Ser Ser Pro
            340                 345                 350
Pro Ala Asp Ser Glu Gly Asn Thr Lys Asp Thr Asn Cys Leu Ser Arg
        355                 360                 365
Ala Glu Gly Pro Gly Ala Gln Asp Gly Trp Val Cys Glu His Arg Trp
    370                 375                 380
Pro Met Ile Arg Glu Met Val Lys Phe Arg Asn Val Thr Ser Ser Lys
385                 390                 395                 400
Phe Tyr Leu Ser Asn Trp Trp Ser Asn Gly Asn Asn Gln Ile Ala Phe
                405                 410                 415
Ser Arg Gly Asp Leu Gly Phe Val Val Ile Asn His Glu Glu Gln Val
            420                 425                 430
Leu Arg Gln Arg Leu Gln Thr Gly Leu Thr Ala Gly Val Tyr Cys Asp
        435                 440                 445
Val Leu Ser Gln Asn Cys Glu Arg Thr Ile Thr Val Asn Asp Asn Gly
    450                 455                 460
Glu Ala Glu Ile Ile Val Thr Gly Gln Asn Ala Ala Ala Ile His Ala
465                 470                 475                 480
Gln Ser Lys Leu Lys
                485
```

<210> 388
<211> 1662
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 388

```
atgagccaca tcctgcgtgc cgccgtattg gcggcggtcc tgctgccgtt tcccgcactg      60
gccgatcagg ccggcaagag cccggccggg gtgcgctacc acggcggcga cgaaatcatc     120
ctccagggct tccactggaa cgtcgtccgc gaagcgccca acgactggta caacatcctc     180
cgccaacagg cctcgacgat cgcggccgac ggcttctcgg caatctggat gccggtgccc     240
tggcgtgact tctccagctg gagcgacggc ggcaagtccg gcggcggcga aggctacttc     300
tggcacgatt tcaacaagaa cggccgctat ggcagcgatg cccagctgcg ccaggccgcc     360
ggcgcgctcg gcggtgccgg ggtgaaggtg ctctacgacg tggtgcccaa ccacatgaac     420
cgcggctacc cgaacaagga aatcaacctg ccggccggcc agggcttctg gcgcaacgac     480
tgcgccgacc cgggcaacta ccccaacgac tgcgacgacg gcgatcgttt cgtcggcggc     540
gatgccgacc tcaataccgg ccatccgcag gtctatggca tgttccgcga cgagttcgcc     600
aacctgcgca gccagtacgg cgccggcggc ttccgcttcg atttcgtccg cggcttcgcc     660
ccggagcggg tcaacagctg gatgaccgac agtgccgaca acagcttctg cgtcggcgag     720
ctgtggaaag gcccttctga atatcccagc tgggactggc gcaacacggc gagctggcag     780
cagatcatca aggactggtc cgatcgggcc aagtgcccgg tgttcgactt cgccctcaag     840
gagcgcatgc agaacggctc gatcgccgac tggaagaacg gccttaacgg caaccccgat     900
ccgcgctggc gcgaggtggc agtgaccttt gtcgacaacc acgacaccgg ctactcgccc     960
gggcagaacg gtgggcagca tcactggccg ctgcaggacg ggctgatccg tcaggcctac    1020
gcgtacattc tgaccagccc cggcacgccg gtggtgtact ggtcgcacat gtacgattgg    1080
ggctaccgcg acttcatccg ccagctgatc caggtgcgtc gcgccgctgg cgtgcgtgcc    1140
```

gattcggcga tcagcttcca cagcggctac agcggcctcg tcgccaccgt caccggcagt 1200
cagcagaccc tggtggtggc gctcaattcc aacctgagca atcccggcca ggtcgccagc 1260
ggcagcttca gcgaagcagt caataccagt aatggccagg tgcgggtctg gcggaccggc 1320
gccggcagtg gtggcggtga caatggcggc ggtgagcccg gtgccctggt cagcgtgaac 1380
ttccgctgcg acaacggtgt gacgcaaccc ggcgacagcg tctacgcggt gggcaacgtc 1440
agccagctcg gcaactggag cccggcctct gcggtgcggc tgaccgatac cagcggctac 1500
ccgacctgga agggccgcat ctcgctgccg gccggccaga acgtggaatg gaagtgcctg 1560
atccgcaacg aggccaatgc aacgcaggtg cggcaatggc agggcggggc caacaacagc 1620
gtgaagccca ccgaaggtgc caacacggcg gggcggttct ag 1662

<210> 389
<211> 553
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(21)

<400> 389
Met Ser His Ile Leu Arg Ala Ala Val Leu Ala Ala Val Leu Leu Pro
 1   5   10   15
Phe Pro Ala Leu Ala Asp Gln Ala Gly Lys Ser Pro Ala Gly Val Arg
            20   25   30
Tyr His Gly Gly Asp Glu Ile Ile Leu Gln Gly Phe His Trp Asn Val
        35   40   45
Val Arg Glu Ala Pro Asn Asp Trp Tyr Asn Ile Leu Arg Gln Gln Ala
    50   55   60
Ser Thr Ile Ala Ala Asp Gly Phe Ser Ala Ile Trp Met Pro Val Pro
65   70   75   80
Trp Arg Asp Phe Ser Ser Trp Ser Asp Gly Gly Lys Ser Gly Gly Gly
                85   90   95
Glu Gly Tyr Phe Trp His Asp Phe Asn Lys Asn Gly Arg Tyr Gly Ser
            100   105   110
Asp Ala Gln Leu Arg Gln Ala Ala Gly Ala Leu Gly Gly Ala Gly Val
        115   120   125
Lys Val Leu Tyr Asp Val Val Pro Asn His Met Asn Arg Gly Tyr Pro
    130   135   140
Asn Lys Glu Ile Asn Leu Pro Ala Gly Gln Gly Phe Trp Arg Asn Asp
145   150   155   160
Cys Ala Asp Pro Gly Asn Tyr Pro Asn Asp Cys Asp Asp Gly Asp Arg
                165   170   175
Phe Val Gly Gly Asp Ala Asp Leu Asn Thr Gly His Pro Gln Val Tyr
            180   185   190
Gly Met Phe Arg Asp Glu Phe Ala Asn Leu Arg Ser Gln Tyr Gly Ala
        195   200   205
Gly Gly Phe Arg Phe Asp Phe Val Arg Gly Phe Ala Pro Glu Arg Val
    210   215   220
Asn Ser Trp Met Thr Asp Ser Ala Asp Asn Ser Phe Cys Val Gly Glu
225   230   235   240
Leu Trp Lys Gly Pro Ser Glu Tyr Pro Ser Trp Asp Trp Arg Asn Thr
                245   250   255
Ala Ser Trp Gln Gln Ile Ile Lys Asp Trp Ser Asp Arg Ala Lys Cys
            260   265   270
Pro Val Phe Asp Phe Ala Leu Lys Glu Arg Met Gln Asn Gly Ser Ile
        275   280   285
Ala Asp Trp Lys Asn Gly Leu Asn Gly Asn Pro Asp Pro Arg Trp Arg
    290   295   300
Glu Val Ala Val Thr Phe Val Asp Asn His Asp Thr Gly Tyr Ser Pro
305   310   315   320
Gly Gln Asn Gly Gly Gln His His Trp Pro Leu Gln Asp Gly Leu Ile
                325   330   335
Arg Gln Ala Tyr Ala Tyr Ile Leu Thr Ser Pro Gly Thr Pro Val Val
            340   345   350
Tyr Trp Ser His Met Tyr Asp Trp Gly Tyr Arg Asp Phe Ile Arg Gln
        355   360   365

Leu Ile Gln Val Arg Arg Ala Ala Gly Val Arg Ala Asp Ser Ala Ile
370 375 380
Ser Phe His Ser Gly Tyr Ser Gly Leu Val Ala Thr Val Thr Gly Ser
385 390 395 400
Gln Gln Thr Leu Val Val Ala Leu Asn Ser Asn Leu Ser Asn Pro Gly
405 410 415
Gln Val Ala Ser Gly Ser Phe Ser Glu Ala Val Asn Thr Ser Asn Gly
420 425 430
Gln Val Arg Val Trp Arg Thr Gly Ala Gly Ser Gly Gly Gly Asp Asn
435 440 445
Gly Gly Gly Glu Pro Gly Ala Leu Val Ser Val Asn Phe Arg Cys Asp
450 455 460
Asn Gly Val Thr Gln Pro Gly Asp Ser Val Tyr Ala Val Gly Asn Val
465 470 475 480
Ser Gln Leu Gly Asn Trp Ser Pro Ala Ser Ala Val Arg Leu Thr Asp
485 490 495
Thr Ser Gly Tyr Pro Thr Trp Lys Gly Arg Ile Ser Leu Pro Ala Gly
500 505 510
Gln Asn Val Glu Trp Lys Cys Leu Ile Arg Asn Glu Ala Asn Ala Thr
515 520 525
Gln Val Arg Gln Trp Gln Gly Gly Ala Asn Asn Ser Val Lys Pro Thr
530 535 540
Glu Gly Ala Asn Thr Ala Gly Arg Phe
545 550

<210> 390
<211> 1497
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 390
atgaaccgtc actcgacggg tgcctcgacg cacccgcagc agcctcgttt tgcccgccgc 60
ctggcggagt gcatccacac acgtctggct tccgccatcc gcagcgccgc ccgggccctg 120
cccggcatcg ccgcgcttgc agccttcgcc ggcagcacga cgccggcccg ggccgacgtc 180
atcctgcacg ccttcaactg gcgctacgac actgtcgaag cacgcgccgc cgagatccag 240
tcgctgggct acaaggcggt gctggtggcg ccgccgctga agtcggaggg ctcggcctgg 300
tgggcgcgct accagccgca ggactaccgc acgatcgacc atccgctggg caacaagcag 360
agcttccagc gcatgtcgaa cgcgctgcgc gcacgcggca tccgcgtcta cgccgacatc 420
atcctcaacc acatggccaa tgaggcggcc cagcgcggcg acctcaacta tcccggccag 480
cgcatcctca acgcctacgc tgccaacacc agctactgga acaaccagcg cctgtttggc 540
gacctgcgct acaacttcct gtcgcagtgg gacttcggcc cggccaactg catcagcaac 600
tacaacgacg tctggcaggt gcagaactgg cgcctctgcg gcagcccagg cgacgccggc 660
ctgcccgacc tgctcggcag cagctatgtg gtcgggcagc agcgcagcta cctgacggca 720
ctcaaggggc tcggcgtcac cggcttccgc atcgatgccg ccaagcacat gccgctctcg 780
cacatcaacg cggtgctgac acccgagatc aagcagggcg tgcatgtgtt cggcgaggtg 840
atcaccaacg gcggcgccgg cgacagcgag tacaacggct tcctcgcccc ctacttgaac 900
ggcaccgacc acgcggccta cgactttccg ctgttcggca gcatccgcaa cgccttccgc 960
ttcggcggca gcatgaacgc gctggtcaac ccgctggccg tcggccaggc cctgcccaac 1020
agccgggctg tgacgttcac cgtgacccac gacatcccga acaacggcgg cttccgctac 1080
ctgatcctcg accccaccga cgagaccctg gcctatgcct atgtgctggg ccgcgacggc 1140
ggcagcccgc tgctgtactc ggacaacaac gagagcggcg acaaccgctg ggtcaacgcc 1200
taccgccgca gtgacctcgc cgcgatgatc cgcttccaca acgccgccca gggcagcgac 1260
atgcaggtgc tgtcctacag cgactgccac ctgctgttcc gccgcggcaa ccgcggcatc 1320
gtcggtatca acaaatgcgg cagcacggtc aacaccacca tcaacatgaa caacagcgtg 1380
ctctggtggt acaccaatta ccgtgatgtt ctggacgcga acagcgtcgt caacatcggc 1440
agcagcagct acaccttcag cctgccggcg cgcagggcga ggatgtggtt gaggtag 1497

<210> 391
<211> 498
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(58)

<400> 391

```
Met Asn Arg His Ser Thr Gly Ala Ser Thr His Pro Gln Gln Pro Arg
 1               5                  10                  15
Phe Ala Arg Arg Leu Ala Glu Cys Ile His Thr Arg Leu Ala Ser Ala
            20                  25                  30
Ile Arg Ser Ala Ala Arg Ala Leu Pro Gly Ile Ala Ala Leu Ala Ala
        35                  40                  45
Phe Ala Gly Ser Thr Thr Pro Ala Arg Ala Asp Val Ile Leu His Ala
    50                  55                  60
Phe Asn Trp Arg Tyr Asp Thr Val Glu Ala Arg Ala Ala Glu Ile Gln
65                  70                  75                  80
Ser Leu Gly Tyr Lys Ala Val Leu Val Ala Pro Pro Leu Lys Ser Glu
                85                  90                  95
Gly Ser Ala Trp Trp Ala Arg Tyr Gln Pro Gln Asp Tyr Arg Thr Ile
            100                 105                 110
Asp His Pro Leu Gly Asn Lys Gln Ser Phe Gln Arg Met Ser Asn Ala
        115                 120                 125
Leu Arg Ala Arg Gly Ile Arg Val Tyr Ala Asp Ile Ile Leu Asn His
    130                 135                 140
Met Ala Asn Glu Ala Ala Gln Arg Gly Asp Leu Asn Tyr Pro Gly Gln
145                 150                 155                 160
Arg Ile Leu Asn Ala Tyr Ala Ala Asn Thr Ser Tyr Trp Asn Asn Gln
                165                 170                 175
Arg Leu Phe Gly Asp Leu Arg Tyr Asn Phe Leu Ser Gln Trp Asp Phe
            180                 185                 190
Gly Pro Ala Asn Cys Ile Ser Asn Tyr Asn Asp Val Trp Gln Val Gln
        195                 200                 205
Asn Trp Arg Leu Cys Gly Ser Pro Gly Asp Ala Gly Leu Pro Asp Leu
    210                 215                 220
Leu Gly Ser Ser Tyr Val Val Gly Gln Gln Arg Ser Tyr Leu Thr Ala
225                 230                 235                 240
Leu Lys Gly Leu Gly Val Thr Gly Phe Arg Ile Asp Ala Ala Lys His
                245                 250                 255
Met Pro Leu Ser His Ile Asn Ala Val Leu Thr Pro Glu Ile Lys Gln
            260                 265                 270
Gly Val His Val Phe Gly Glu Val Ile Thr Asn Gly Gly Ala Gly Asp
        275                 280                 285
Ser Glu Tyr Asn Gly Phe Leu Ala Pro Tyr Leu Asn Gly Thr Asp His
    290                 295                 300
Ala Ala Tyr Asp Phe Pro Leu Phe Gly Ser Ile Arg Asn Ala Phe Arg
305                 310                 315                 320
Phe Gly Gly Ser Met Asn Ala Leu Val Asn Pro Leu Ala Val Gly Gln
                325                 330                 335
Ala Leu Pro Asn Ser Arg Ala Val Thr Phe Thr Val Thr His Asp Ile
            340                 345                 350
Pro Asn Asn Gly Gly Phe Arg Tyr Leu Ile Leu Asp Pro Thr Asp Glu
        355                 360                 365
Thr Leu Ala Tyr Ala Tyr Val Leu Gly Arg Asp Gly Gly Ser Pro Leu
    370                 375                 380
Leu Tyr Ser Asp Asn Asn Glu Ser Gly Asp Asn Arg Trp Val Asn Ala
385                 390                 395                 400
Tyr Arg Arg Ser Asp Leu Ala Ala Met Ile Arg Phe His Asn Ala Ala
                405                 410                 415
Gln Gly Ser Asp Met Gln Val Leu Ser Tyr Ser Asp Cys His Leu Leu
            420                 425                 430
Phe Arg Arg Gly Asn Arg Gly Ile Val Gly Ile Asn Lys Cys Gly Ser
        435                 440                 445
Thr Val Asn Thr Thr Ile Asn Met Asn Asn Ser Val Leu Trp Trp Tyr
    450                 455                 460
Thr Asn Tyr Arg Asp Val Leu Asp Ala Asn Ser Val Val Asn Ile Gly
465                 470                 475                 480
Ser Ser Ser Tyr Thr Phe Ser Leu Pro Ala Arg Arg Ala Arg Met Trp
                485                 490                 495
Leu Arg
```

<210> 392
<211> 2100
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 392
atgccttcag ggctcaagca tggcatcaat atccacgatg acaattctgt cacattcgtc 60
ctatatgctc aggacaagga caaggccatg catgagtatt gctacctcgt aggtgattgg 120
aacaactggg agcgcactgc cgagagttcg atgaagaggg atacggaggc cggttgctgg 180
tggctcaagg tcggaggatt tgatccggca aaggagtata gatataattt ccgtctcggc 240
aacccgcctg catcagacat ctgtgtcagc gatccatata caagaatcgt ctatgatcag 300
tggaatgacc agtatcttga aggtgttcct gcattccctg agggcgcgaa ggctctcgtt 360
tctgcattcc agataagcaa accggcctac tcatggaagg ttacggactt caagatagag 420
gacaggaagg atctcgtgat ctatgagatg ctgttccgcg atttctcagc gacaaaggat 480
atcgcaggag cggtcgccgg gtttgatcac atcaagaacc tcggtgtgaa tgcagtccag 540
ctcatgccta ttcaggaatt tgacggcaac aacagttggg gatataaccc taaccactat 600
ttcgcccttg acaaggctta tggaacgctt gacgagtaca agcagttcat tgatctttgt 660
catgaaaacg gcattgctgt gatcgtggat gtcgtctata accacacctt cggttctcat 720
acatgggcaa agatgtggtg ggattccgca aacaacagaa ctgcttccaa caatccttgg 780
tacaatgaat tcccgacaca tccgtacaat gtgggccatg acttcaatca tgagaatgag 840
atggttgacc agcatgtgaa ggaaagtctt gaattccttc tcagggagta caaggtggac 900
ggcttccgtt tcgacctcac caagggtttc acccagaaga agacagaccc ggatgttgcg 960
gcatggggac gttacgacca gagccgtgtg gatatcctca aggactatgc ggaccatatc 1020
tggtcggtga agaatgatgc tgtagttatc ttcgagcatc tttctgactg ggacgaagag 1080
gaagttctcg caaatcatgg aatccagctt tggagaaatg taaataaaga atatagttcg 1140
gctgtcaccg gcggaacagg tgatttcaag aatatgcact ctacctcacc gttcggcgga 1200
tatgtgggat atatggagag ccacgatgag gagcgtatat gctatggcgc ttctgacgat 1260
gtatcttcga tcacatgggg tgtcatcgga cttggaaatg attgggacaa cgacaaggag 1320
atgacggcag agaacggact gttcgttgta aagaatgtta cggttgccgc caatgacaag 1380
ttcaagatcc gcaaggccaa ggaatggaat gacaactaca actacggtgc tgccaatgat 1440
aacttcaaac tgacagtcgg tcagggatac aagatgacca atggaaacag ttcaaaggac 1500
atgggcgtcc ctgttgcagg caagtatgac atctggttca gccatgatat tgccaccgta 1560
tggcttatgc ctgccggcca gaagccggat tatactccgt cgacttccgt atcatctgag 1620
gatgccctca ctgtagccat gcgtcgtgca ggtgcatctg ccgccttctt cctcacagtg 1680
cctggaccta agatgatctg gcagttcggt gagatcggat atgatctttc gatcaactat 1740
ccttccggca ctgaagatga caggacaagc gaaaagccgg taaagaccgc tgaatacatg 1800
aagaatcctg cacgcaagtc tctctacgac acctatgcag gtctgttgaa gttccgccgc 1860
gagaaccctc gcttctttga ttctgatgcg aagttcgagt ggactccttc cggtgagatc 1920
aagaagatca cctgctctgt ggacggcaag accttccatg tagtgggtaa cttctcgcag 1980
atggacaagt catattcagt tccgtcaggc agttggaagg attacttcaa tggcggagaa 2040
tctgtttcgg gaaccattac cctcaagcag ggtgagttca ggctcctgac cagcttctga 2100

<210> 393
<211> 699
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 393
Met Pro Ser Gly Leu Lys His Gly Ile Asn Ile His Asp Asp Asn Ser
1 5 10 15
Val Thr Phe Val Leu Tyr Ala Gln Asp Lys Asp Lys Ala Met His Glu
20 25 30
Tyr Cys Tyr Leu Val Gly Asp Trp Asn Asn Trp Glu Arg Thr Ala Glu
35 40 45
Ser Ser Met Lys Arg Asp Thr Glu Ala Gly Cys Trp Trp Leu Lys Val
50 55 60
Gly Gly Phe Asp Pro Ala Lys Glu Tyr Arg Tyr Asn Phe Arg Leu Gly
65 70 75 80
Asn Pro Pro Ala Ser Asp Ile Cys Val Ser Asp Pro Tyr Thr Arg Ile
85 90 95
Val Tyr Asp Gln Trp Asn Asp Gln Tyr Leu Glu Gly Val Pro Ala Phe

```
            100                 105                 110
Pro Glu Gly Ala Lys Ala Leu Val Ser Ala Phe Gln Ile Ser Lys Pro
        115                 120                 125
Ala Tyr Ser Trp Lys Val Thr Asp Phe Lys Ile Glu Asp Arg Lys Asp
    130                 135                 140
Leu Val Ile Tyr Glu Met Leu Phe Arg Asp Phe Ser Ala Thr Lys Asp
145                 150                 155                 160
Ile Ala Gly Ala Val Ala Gly Phe Asp His Ile Lys Asn Leu Gly Val
                165                 170                 175
Asn Ala Val Gln Leu Met Pro Ile Gln Glu Phe Asp Gly Asn Asn Ser
            180                 185                 190
Trp Gly Tyr Asn Pro Asn His Tyr Phe Ala Leu Asp Lys Ala Tyr Gly
        195                 200                 205
Thr Leu Asp Glu Tyr Lys Gln Phe Ile Asp Leu Cys His Glu Asn Gly
    210                 215                 220
Ile Ala Val Ile Val Asp Val Val Tyr Asn His Thr Phe Gly Ser His
225                 230                 235                 240
Thr Trp Ala Lys Met Trp Trp Asp Ser Ala Asn Asn Arg Thr Ala Ser
                245                 250                 255
Asn Asn Pro Trp Tyr Asn Glu Phe Pro Thr His Pro Tyr Asn Val Gly
            260                 265                 270
His Asp Phe Asn His Glu Asn Glu Met Val Asp Gln His Val Lys Glu
        275                 280                 285
Ser Leu Glu Phe Leu Leu Arg Glu Tyr Lys Val Asp Gly Phe Arg Phe
    290                 295                 300
Asp Leu Thr Lys Gly Phe Thr Gln Lys Lys Thr Asp Pro Asp Val Ala
305                 310                 315                 320
Ala Trp Gly Arg Tyr Asp Gln Ser Arg Val Asp Ile Leu Lys Asp Tyr
                325                 330                 335
Ala Asp His Ile Trp Ser Val Lys Asn Asp Ala Val Val Ile Phe Glu
            340                 345                 350
His Leu Ser Asp Trp Asp Glu Glu Glu Val Leu Ala Asn His Gly Ile
        355                 360                 365
Gln Leu Trp Arg Asn Val Asn Lys Glu Tyr Ser Ser Ala Val Thr Gly
    370                 375                 380
Gly Thr Gly Asp Phe Lys Asn Met His Ser Thr Ser Pro Phe Gly Gly
385                 390                 395                 400
Tyr Val Gly Tyr Met Glu Ser His Asp Glu Glu Arg Ile Cys Tyr Gly
                405                 410                 415
Ala Ser Asp Asp Val Ser Ser Ile Thr Trp Gly Val Ile Gly Leu Gly
            420                 425                 430
Asn Asp Trp Asp Asn Asp Lys Glu Met Thr Ala Glu Asn Gly Leu Phe
        435                 440                 445
Val Val Lys Asn Val Thr Val Ala Ala Asn Asp Lys Phe Lys Ile Arg
    450                 455                 460
Lys Ala Lys Glu Trp Asn Asp Asn Tyr Asn Tyr Gly Ala Ala Asn Asp
465                 470                 475                 480
Asn Phe Lys Leu Thr Val Gly Gln Gly Tyr Lys Met Thr Asn Gly Asn
                485                 490                 495
Ser Ser Lys Asp Met Gly Val Pro Val Ala Gly Lys Tyr Asp Ile Trp
            500                 505                 510
Phe Ser His Asp Ile Ala Thr Val Trp Leu Met Pro Ala Gly Gln Lys
        515                 520                 525
Pro Asp Tyr Thr Pro Ser Thr Ser Val Ser Ser Glu Asp Ala Leu Thr
    530                 535                 540
Val Ala Met Arg Arg Ala Gly Ala Ser Ala Ala Phe Phe Leu Thr Val
545                 550                 555                 560
Pro Gly Pro Lys Met Ile Trp Gln Phe Gly Glu Ile Gly Tyr Asp Leu
                565                 570                 575
Ser Ile Asn Tyr Pro Ser Gly Thr Glu Asp Asp Arg Thr Ser Glu Lys
            580                 585                 590
Pro Val Lys Thr Ala Glu Tyr Met Lys Asn Pro Ala Arg Lys Ser Leu
        595                 600                 605
Tyr Asp Thr Tyr Ala Gly Leu Leu Lys Phe Arg Arg Glu Asn Pro Arg
    610                 615                 620
Phe Phe Asp Ser Asp Ala Lys Phe Glu Trp Thr Pro Ser Gly Glu Ile
625                 630                 635                 640
Lys Lys Ile Thr Cys Ser Val Asp Gly Lys Thr Phe His Val Val Gly
```

645 650 655
Asn Phe Ser Gln Met Asp Lys Ser Tyr Ser Val Pro Ser Gly Ser Trp
660 665 670
Lys Asp Tyr Phe Asn Gly Gly Glu Ser Val Ser Gly Thr Ile Thr Leu
675 680 685
Lys Gln Gly Glu Phe Arg Leu Leu Thr Ser Phe
690 695

<210> 394
<211> 1347
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 394
atgcggacat cagtgttgtg cggagtggtc ctgatccaac aggcccgtcg tttcgcttcg 60
agaaaatcag ttcatatcca cgccagcagc acctgcagtc ggttgcggcg ctggcacatc 120
gtcgcctcgc tgctcgccag tacctggctg gcgaccgagc cggcgcaagc ggccacggaa 180
cagaatcaag ccagcagcgc gatactgttc cagggttttc actggcactc cgccaacggt 240
agctggtacg gcaatctgca gggcaaggct ggcgatttga aagacttggg catcactcac 300
gtctggttcc cgccgccgtc ggacgctgcg tcactggaag ggtatctgcc gcgccagctc 360
aatgtgctga acagcagtta cggcagtgag acggcgctac gcaatgccat cagcgcgctg 420
accggccaag gcatcaagag cgtggccgat gtcgtgatca atcaccgcgt tggcaccagc 480
aactgggccg acttcaccaa tccgagctgg gattgccgcg ctgtggtcaa caatgacgaa 540
tggagcgggc gttgcggcaa cggcgattcc ggtgacggct atggcgctgc ccgcgatctg 600
gatcactcgc aaggctttgt tcagaacgat ctgaaaacct ggctgtcgtc gcgactcaaa 660
ggcgttggtt tttccggtat tcgttacgac tactcgaaag gctacggttc cggttacgcc 720
ggcctgtatc acgacgcaat gaatccgaac ttctgcgtcg gtgagatctg gaccgatctg 780
gattacaaca atgtcgatgc gcatcgccag ctgttgatga attatgtcga cggcaacaat 840
ggcaaatgcg gcgccttcga tttcaccagc aaaggtctgc tcaatcaggc actgtgggcc 900
aacgaatatt ggcgactgcg cgcctgggat ggcaaaccgg ctggtggcat cggttggtgg 960
gcgcagaaga tggtgacgtt cgtcgacaac catgacaccg gtccgtcgca gagttgcagt 1020
gccggccaga accactggcc ggtgccgtgc ggcaaagtga tgcagggtta tgcctacatc 1080
ctgagtcatc ccggtattcc gaccgtgtat tacccgcacg tctatgactg gaacctgcgc 1140
agtgccatca agaccctgat cagcctgcgc aaggagaagg gcatcacctc gaccagcagc 1200
gtcgcgatcc agcgcgccga tcaaggcctc tatgccgcca tcatcaacgg caatctggcg 1260
atgaagatcg gcccgaacag ctggagtccc ggcagcggct ggacgctgcg catgtctggc 1320
gaacagtacg caatttggac ccgttga 1347

<210> 395
<211> 448
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(57)

<400> 395
Met Arg Thr Ser Val Leu Cys Gly Val Val Leu Ile Gln Gln Ala Arg
1 5 10 15
Arg Phe Ala Ser Arg Lys Ser Val His Ile His Ala Ser Ser Thr Cys
20 25 30
Ser Arg Leu Arg Arg Trp His Ile Val Ala Ser Leu Leu Ala Ser Thr
35 40 45
Trp Leu Ala Thr Glu Pro Ala Gln Ala Ala Thr Glu Gln Asn Gln Ala
50 55 60
Ser Ser Ala Ile Leu Phe Gln Gly Phe His Trp His Ser Ala Asn Gly
65 70 75 80
Ser Trp Tyr Gly Asn Leu Gln Gly Lys Ala Gly Asp Leu Lys Asp Leu
85 90 95
Gly Ile Thr His Val Trp Phe Pro Pro Pro Ser Asp Ala Ala Ser Leu
100 105 110
Glu Gly Tyr Leu Pro Arg Gln Leu Asn Val Leu Asn Ser Ser Tyr Gly

```
        115                 120                 125
Ser Glu Thr Ala Leu Arg Asn Ala Ile Ser Ala Leu Thr Gly Gln Gly
    130                 135                 140
Ile Lys Ser Val Ala Asp Val Val Ile Asn His Arg Val Gly Thr Ser
145                 150                 155                 160
Asn Trp Ala Asp Phe Thr Asn Pro Ser Trp Asp Cys Arg Ala Val Val
                165                 170                 175
Asn Asn Asp Glu Trp Ser Gly Arg Cys Gly Asn Gly Asp Ser Gly Asp
            180                 185                 190
Gly Tyr Gly Ala Ala Arg Asp Leu Asp His Ser Gln Gly Phe Val Gln
        195                 200                 205
Asn Asp Leu Lys Thr Trp Leu Ser Ser Arg Leu Lys Gly Val Gly Phe
    210                 215                 220
Ser Gly Ile Arg Tyr Asp Tyr Ser Lys Gly Tyr Gly Ser Gly Tyr Ala
225                 230                 235                 240
Gly Leu Tyr His Asp Ala Met Asn Pro Asn Phe Cys Val Gly Glu Ile
                245                 250                 255
Trp Thr Asp Leu Asp Tyr Asn Asn Val Asp Ala His Arg Gln Leu Leu
            260                 265                 270
Met Asn Tyr Val Asp Gly Asn Asn Gly Lys Cys Gly Ala Phe Asp Phe
        275                 280                 285
Thr Ser Lys Gly Leu Leu Asn Gln Ala Leu Trp Ala Asn Glu Tyr Trp
    290                 295                 300
Arg Leu Arg Ala Trp Asp Gly Lys Pro Ala Gly Gly Ile Gly Trp Trp
305                 310                 315                 320
Ala Gln Lys Met Val Thr Phe Val Asp Asn His Asp Thr Gly Pro Ser
                325                 330                 335
Gln Ser Cys Ser Ala Gly Gln Asn His Trp Pro Val Pro Cys Gly Lys
            340                 345                 350
Val Met Gln Gly Tyr Ala Tyr Ile Leu Ser His Pro Gly Ile Pro Thr
        355                 360                 365
Val Tyr Tyr Pro His Val Tyr Asp Trp Asn Leu Arg Ser Ala Ile Lys
    370                 375                 380
Thr Leu Ile Ser Leu Arg Lys Glu Lys Gly Ile Thr Ser Thr Ser Ser
385                 390                 395                 400
Val Ala Ile Gln Arg Ala Asp Gln Gly Leu Tyr Ala Ala Ile Ile Asn
                405                 410                 415
Gly Asn Leu Ala Met Lys Ile Gly Pro Asn Ser Trp Ser Pro Gly Ser
            420                 425                 430
Gly Trp Thr Leu Arg Met Ser Gly Glu Gln Tyr Ala Ile Trp Thr Arg
        435                 440                 445
```

<210> 396
<211> 1644
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 396

```
atgaaaaaaa ccatcttatt aaccaccatt tgcatcctat tatggaattg tcaatccaaa      60
aaaaacatga ccgaacaagc acctttttgat tggaaagcag ccaatgtgta ttttttactc     120
accgatcgct ttgcaaacgg agaccctagc aatgacaacc tcatcaaaag agatttgcca     180
acaggggttt tacgtggttt tatgggagga gatttaaaag gcatcaccca aaaaatagaa     240
gaaggttatt ttaccaactt aggcatcaat gccatttgga tgactccgat tgtagaacaa     300
attcacggtg cagtggatga aggcaccgga aacacctatg gttttcatgg ctattggacc     360
aaagattgga cagctattga ccctaattat ggcaccaaag ccgatttaaa agaattggta     420
gaaaaagccc atgccaaagg cattcgaatt gtattagatg cagtaatcaa ccacactggt     480
ccggttaccg aattagatcc ggtgttccca gatagttggg ttagaaccaa acctaaatgt     540
acttacgaca cgtataaaaa ttatatcgat tgtacgttag tagaaaatct tcccgacatc     600
aaaaccgaaa gcaatgaaga agtagaattg ccaccaatgt tagtggaaaa gtggaaaaaa     660
gaaggcagat atgaagaaga aataagagaa ttagatgctt tttttgccaa aaccggttat     720
ccaagagcac caaaatacta catcatgaaa tggttggcag attacatcac cgactttggt     780
attgatggtt acagagtgga tacggttaaa cacaccacag aagatgtttg ggcagatttt     840
gcaaaagttt gtaaagatgc ttttacagct tataaaactg caaatcctaa aaaagtgcta     900
gacgataacg acttcttttt ggttggagaa gtatatggtt atggtttgca tggcaaacaa     960
atctatgatt ttggagataa aaaagtgaat tattttgaaa acggattcaa ctttttaatc    1020
```

aattttgatt ttaaaggaga tgccaacaaa gggtatgaag aattgtttag ctattacgat 1080
aaagttttgc aaaacgaaat gggcaactat agtgtgatga attacttaac ttctcatgac 1140
gatggttggc cttttgacaa aaaaagagaa cgtacttttg aagcagctac caagttactg 1200
ctttctcctg gagtggcaca agtatattat ggagatgaaa ctgcaagatc tttagaaatt 1260
gaaggcaccc aaggagatgc cactttgcga agttttatga attgggaaga ggttcaaagt 1320
aatgctaaaa ccaaagcatt attagttcat tatcaaaagt tagggtcatt tagagcaaac 1380
cacccggcag ttggagcagg taaacacgaa aaaattagcg atgcacctta tacttttcc 1440
agaacttatc aaaaaggtga tttcaaagat gtcatagtaa ttactttaga cgcttcaaaa 1500
ggtgccaaaa ccatccaagt aaataatgca tttgcagagg gcagtaaagt tagagatgct 1560
tattcaggaa tgattggaga agtaaaaaat ggtaaagtag agatccattc taacgaaact 1620
atcgtgttgt tagaaaaaat gtaa 1644

<210> 397
<211> 547
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(19)

<400> 397
Met Lys Lys Thr Ile Leu Leu Thr Thr Ile Cys Ile Leu Leu Trp Asn
 1   5   10   15
Cys Gln Ser Lys Lys Asn Met Thr Glu Gln Ala Pro Phe Asp Trp Lys
            20   25   30
Ala Ala Asn Val Tyr Phe Leu Leu Thr Asp Arg Phe Ala Asn Gly Asp
        35   40   45
Pro Ser Asn Asp Asn Leu Ile Lys Arg Asp Leu Pro Thr Gly Val Leu
    50   55   60
Arg Gly Phe Met Gly Gly Asp Leu Lys Gly Ile Thr Gln Lys Ile Glu
65   70   75   80
Glu Gly Tyr Phe Thr Asn Leu Gly Ile Asn Ala Ile Trp Met Thr Pro
                85   90   95
Ile Val Glu Gln Ile His Gly Ala Val Asp Glu Gly Thr Gly Asn Thr
            100   105   110
Tyr Gly Phe His Gly Tyr Trp Thr Lys Asp Trp Thr Ala Ile Asp Pro
        115   120   125
Asn Tyr Gly Thr Lys Ala Asp Leu Lys Glu Leu Val Glu Lys Ala His
    130   135   140
Ala Lys Gly Ile Arg Ile Val Leu Asp Ala Val Ile Asn His Thr Gly
145   150   155   160
Pro Val Thr Glu Leu Asp Pro Val Phe Pro Asp Ser Trp Val Arg Thr
                165   170   175
Lys Pro Lys Cys Thr Tyr Asp Thr Tyr Lys Asn Tyr Ile Asp Cys Thr
            180   185   190
Leu Val Glu Asn Leu Pro Asp Ile Lys Thr Glu Ser Asn Glu Glu Val
        195   200   205
Glu Leu Pro Pro Met Leu Val Glu Lys Trp Lys Lys Glu Gly Arg Tyr
    210   215   220
Glu Glu Glu Ile Arg Glu Leu Asp Ala Phe Phe Ala Lys Thr Gly Tyr
225   230   235   240
Pro Arg Ala Pro Lys Tyr Tyr Ile Met Lys Trp Leu Ala Asp Tyr Ile
                245   250   255
Thr Asp Phe Gly Ile Asp Gly Tyr Arg Val Asp Thr Val Lys His Thr
            260   265   270
Thr Glu Asp Val Trp Ala Asp Phe Ala Lys Val Cys Lys Asp Ala Phe
        275   280   285
Thr Ala Tyr Lys Thr Ala Asn Pro Lys Lys Val Leu Asp Asp Asn Asp
    290   295   300
Phe Phe Leu Val Gly Glu Val Tyr Gly Tyr Gly Leu His Gly Lys Gln
305   310   315   320
Ile Tyr Asp Phe Gly Asp Lys Lys Val Asn Tyr Phe Glu Asn Gly Phe
                325   330   335
Asn Phe Leu Ile Asn Phe Asp Phe Lys Gly Asp Ala Asn Lys Gly Tyr
            340   345   350

Glu Glu Leu Phe Ser Tyr Tyr Asp Lys Val Leu Gln Asn Glu Met Gly
355 360 365
Asn Tyr Ser Val Met Asn Tyr Leu Thr Ser His Asp Asp Gly Trp Pro
370 375 380
Phe Asp Lys Lys Arg Glu Arg Thr Phe Glu Ala Ala Thr Lys Leu Leu
385 390 395 400
Leu Ser Pro Gly Val Ala Gln Val Tyr Tyr Gly Asp Glu Thr Ala Arg
405 410 415
Ser Leu Glu Ile Glu Gly Thr Gln Gly Asp Ala Thr Leu Arg Ser Phe
420 425 430
Met Asn Trp Glu Glu Val Gln Ser Asn Ala Lys Thr Lys Ala Leu Leu
435 440 445
Val His Tyr Gln Lys Leu Gly Ser Phe Arg Ala Asn His Pro Ala Val
450 455 460
Gly Ala Gly Lys His Glu Lys Ile Ser Asp Ala Pro Tyr Thr Phe Ser
465 470 475 480
Arg Thr Tyr Gln Lys Gly Asp Phe Lys Asp Val Ile Val Ile Thr Leu
485 490 495
Asp Ala Ser Lys Gly Ala Lys Thr Ile Gln Val Asn Asn Ala Phe Ala
500 505 510
Glu Gly Ser Lys Val Arg Asp Ala Tyr Ser Gly Met Ile Gly Glu Val
515 520 525
Lys Asn Gly Lys Val Glu Ile His Ser Asn Glu Thr Ile Val Leu Leu
530 535 540
Glu Lys Met
545

<210> 398
<211> 2040
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 398
gtgaccttcg tcttctacga tcaggatacg gcaggaaaga gccacaagta ctgctacatc 60
gtaggtgact ggaacaactg gacacgcgag acagagggag ccatgtacta cgatggaacc 120
aagggctgct ggtggatcac tcttgacggc ttcgacgctg acaaggagta ccgcttccag 180
tatcgtctcg gcaattcctc aggtactgac acctatgtca gcgacccata cacagagatc 240
gtatatgacc agtggaatga ccagtatatt cagggtgttc ctgcattccc tgaaggagca 300
aaggccctcg tttcggcatt ccagatcaat agacctgaat acagctggaa gcacaaggac 360
ttcaaggtgg aggacaagaa cgacctcgtg atctacgaga tgctcttccg cgacttcgct 420
acatctcaga acattgccgg tgcgatgacg cagctcgact atatcgagaa cctcggtgtc 480
actgcagtcc agctcatgcc gatccaggag ttcgacggca acctcagctg gggatataac 540
cctaaccact ggttcgccct cgacaagtac tacggtacac gtgagcagta caaggagttc 600
atcgacgagt gccacgctcg cggaatggct gtgatcatcg atgtcgtata taaccacgcc 660
acaggttcgc atccttgggc gaagatgtgg tgggacggtg acaggactgc cgcgaacaac 720
ccttggttca acgagtacgc aaagcacccg tacaacgtat atcatgacat gaatcatgag 780
aatgcgatgg tcaaggagca tgtgaagaag agcctcgagt atctcctgac ggagtacgac 840
gtggacggat tccgtttcga cctcacgaaa ggcttcaccc agaaggatac cggcggagac 900
aacggcgacg tggcagcgtg gggacgttat gaccagagcc gtgtggacat cctcaaggga 960
tatgccgacc acatctggtc tgtgaacagt gatgctgtag tcatcttcga gcatctttcc 1020
gactgggatg aggagaaagt tctcgctgag catggcatcc agctttggag aaatgtgaac 1080
ggagaatatc gttcagcggt gaccggcgga acaggaaact tcacgaatat gtattccaac 1140
gctccgttcg gcggattcgt gggatatatg gagagtcacg atgaggagcg tatctgctac 1200
ggtgccggtg ctgatgcatc ttctgtgaca tggggtgtga tcggtctcgg tgacgactgg 1260
aacaatgaca aggtaatgtc taaggacggt gtcttcttct cagtgaagaa tgtgaccgtg 1320
actgcgaacg accgtttcaa gatccgcaag gcaggcgagt ggaacgacgc atttaactac 1380
ggtgcatctg gcgacaactt caagctcacc gtgggtcagg gctacaagat gaccaacggt 1440
aacagctcga aggatatgta tgtaccggct gcaggaacct acgatatcta cttcagcctc 1500
gagactgaga cagtatggct catggctgcc ggtcagcgcc ctgccgatcc ttcgacaggc 1560
ggcggagctt ccgaggatgc ccttacggtg gcaatgcgtc gtgcaggtgc ttctgcagca 1620
ttcttcctca ctgttcctgg tccgaagatg atctggcagt tcggtgagat cggatacgac 1680
tattctatcg agtacaacga ccgcacaggt gagaagcctg tcgtgaccga ccagtacatg 1740
gcagtccctg cccgcaaggc tctctatgac acctatgcgg ctctcctgaa gttccgtcgt 1800
gacaaccctc gcttcttcga caaggatgcg aagttcgagt ggactccttc aggcgagatc 1860
aagaagatca cctgctctgt ggacggaaag accttccatg tggtgggtaa cttcgcgaag 1920

tcttcaaaga catacaccgt tccttcaggc agctggaagg actacttcaa taacggagca 1980
tcggtttcag gaaccatcac cctcaagcag ggcgagttca gactcctgac cagcttctaa 2040

<210> 399
<211> 679
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 399
Met Thr Phe Val Phe Tyr Asp Gln Asp Thr Ala Gly Lys Ser His Lys
 1   5   10   15
Tyr Cys Tyr Ile Val Gly Asp Trp Asn Asn Trp Thr Arg Glu Thr Glu
            20   25   30
Gly Ala Met Tyr Tyr Asp Gly Thr Lys Gly Cys Trp Trp Ile Thr Leu
        35   40   45
Asp Gly Phe Asp Ala Asp Lys Glu Tyr Arg Phe Gln Tyr Arg Leu Gly
    50   55   60
Asn Ser Ser Gly Thr Asp Thr Tyr Val Ser Asp Pro Tyr Thr Glu Ile
65   70   75   80
Val Tyr Asp Gln Trp Asn Asp Gln Tyr Ile Gln Gly Val Pro Ala Phe
                85   90   95
Pro Glu Gly Ala Lys Ala Leu Val Ser Ala Phe Gln Ile Asn Arg Pro
            100   105   110
Glu Tyr Ser Trp Lys His Lys Asp Phe Lys Val Glu Asp Lys Asn Asp
        115   120   125
Leu Val Ile Tyr Glu Met Leu Phe Arg Asp Phe Ala Thr Ser Gln Asn
    130   135   140
Ile Ala Gly Ala Met Thr Gln Leu Asp Tyr Ile Glu Asn Leu Gly Val
145   150   155   160
Thr Ala Val Gln Leu Met Pro Ile Gln Glu Phe Asp Gly Asn Leu Ser
                165   170   175
Trp Gly Tyr Asn Pro Asn His Trp Phe Ala Leu Asp Lys Tyr Tyr Gly
            180   185   190
Thr Arg Glu Gln Tyr Lys Glu Phe Ile Asp Glu Cys His Ala Arg Gly
        195   200   205
Met Ala Val Ile Ile Asp Val Val Tyr Asn His Ala Thr Gly Ser His
    210   215   220
Pro Trp Ala Lys Met Trp Trp Asp Gly Asp Arg Thr Ala Ala Asn Asn
225   230   235   240
Pro Trp Phe Asn Glu Tyr Ala Lys His Pro Tyr Asn Val Tyr His Asp
                245   250   255
Met Asn His Glu Asn Ala Met Val Lys Glu His Val Lys Lys Ser Leu
            260   265   270
Glu Tyr Leu Leu Thr Glu Tyr Asp Val Asp Gly Phe Arg Phe Asp Leu
        275   280   285
Thr Lys Gly Phe Thr Gln Lys Asp Thr Gly Gly Asp Asn Gly Asp Val
    290   295   300
Ala Ala Trp Gly Arg Tyr Asp Gln Ser Arg Val Asp Ile Leu Lys Gly
305   310   315   320
Tyr Ala Asp His Ile Trp Ser Val Asn Ser Asp Ala Val Val Ile Phe
                325   330   335
Glu His Leu Ser Asp Trp Asp Glu Glu Lys Val Leu Ala Glu His Gly
            340   345   350
Ile Gln Leu Trp Arg Asn Val Asn Gly Glu Tyr Arg Ser Ala Val Thr
        355   360   365
Gly Gly Thr Gly Asn Phe Thr Asn Met Tyr Ser Asn Ala Pro Phe Gly
    370   375   380
Gly Phe Val Gly Tyr Met Glu Ser His Asp Glu Glu Arg Ile Cys Tyr
385   390   395   400
Gly Ala Gly Ala Asp Ala Ser Ser Val Thr Trp Gly Val Ile Gly Leu
                405   410   415
Gly Asp Asp Trp Asn Asn Asp Lys Val Met Ser Lys Asp Gly Val Phe
            420   425   430
Phe Ser Val Lys Asn Val Thr Val Thr Ala Asn Asp Arg Phe Lys Ile
        435   440   445

```
Arg Lys Ala Gly Glu Trp Asn Asp Ala Phe Asn Tyr Gly Ala Ser Gly
    450                 455                 460
Asp Asn Phe Lys Leu Thr Val Gly Gln Gly Tyr Lys Met Thr Asn Gly
465                 470                 475                 480
Asn Ser Ser Lys Asp Met Tyr Val Pro Ala Ala Gly Thr Tyr Asp Ile
                485                 490                 495
Tyr Phe Ser Leu Glu Thr Glu Thr Val Trp Leu Met Ala Ala Gly Gln
            500                 505                 510
Arg Pro Ala Asp Pro Ser Thr Gly Gly Gly Ala Ser Glu Asp Ala Leu
        515                 520                 525
Thr Val Ala Met Arg Arg Ala Gly Ala Ser Ala Ala Phe Phe Leu Thr
    530                 535                 540
Val Pro Gly Pro Lys Met Ile Trp Gln Phe Gly Glu Ile Gly Tyr Asp
545                 550                 555                 560
Tyr Ser Ile Glu Tyr Asn Asp Arg Thr Gly Glu Lys Pro Val Val Thr
                565                 570                 575
Asp Gln Tyr Met Ala Val Pro Ala Arg Lys Ala Leu Tyr Asp Thr Tyr
            580                 585                 590
Ala Ala Leu Leu Lys Phe Arg Arg Asp Asn Pro Arg Phe Phe Asp Lys
        595                 600                 605
Asp Ala Lys Phe Glu Trp Thr Pro Ser Gly Glu Ile Lys Lys Ile Thr
    610                 615                 620
Cys Ser Val Asp Gly Lys Thr Phe His Val Val Gly Asn Phe Ala Lys
625                 630                 635                 640
Ser Ser Lys Thr Tyr Thr Val Pro Ser Gly Ser Trp Lys Asp Tyr Phe
                645                 650                 655
Asn Asn Gly Ala Ser Val Ser Gly Thr Ile Thr Leu Lys Gln Gly Glu
            660                 665                 670
Phe Arg Leu Leu Thr Ser Phe
        675
```

<210> 400
<211> 1245
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 400

```
atgaaaaact ttatcttggg aatggcggcc atgctgatgg ccgcgcagag tctggccgcc      60
actgagcaaa acaaagcttc gacagccatc ctgctccagg gctttcactg gaattcggcg     120
aacgggtcct ggtacctgaa cctcgaaggc aaggccgcgg acttccggga tcttggcatc     180
acgcacgtct ggtttccacc gccatcggac tcgggagcgc gcgaaggcta tctgcctcgg     240
gagctttata aacttcagtc ggcctatggc agcgaggagc agctccgcaa tgccatcagc     300
gctctgaatc gggaaggcat ccaggcggtg gccgatatcg taatcaacca tcgcgtcggc     360
agctggaact ggggagactt taactaccct aaatgggact gcgatagcgt cgcgaacaac     420
gatgaatggc ctggccgctg cggtggctat gattcgggtg ccagctatgg agccgcccgg     480
gacatcgacc attccaacgc tcaggtgcag aatgatatca aatatttcct gagcgaacgt     540
ctgcgagggg ttggtttcag cggctggcgc tacgattatt ccaaaggtta tgcaccctat     600
tacgccagga tttataacga agccagccgc cctgactttt gcgtgggcga agtctggccc     660
ggcttgaatt atgatgatgt caacgcgcat cgccaggagc tgatgaatta tgtcgacggc     720
accggaggaa gctgtggagc cttcgacttc acgagcaaag gacttctgaa caaggtcctt     780
cgtgacaatg actactggcg cctgcgcgat gccaatggac gtccggccgg tggcatcggc     840
tggtggccgc agaagatggt gaccttcgtc gataaccatg atactggacc ttccgaatcc     900
tgttcagcag gacagaacat gtggccggtt ccctgtgaca aggtgatgct gggctatgcc     960
tatatcctga ctcatcctgg agtgccgagc ctttattatc cgcatcttta cgattggggt    1020
ttgaagaata gcatcaagtc tctcttgaat gcacgccgct ccgtaggcat tcaatccaac    1080
tccaatatca gcattcaaaa agcagaaaac ggactgtatg cagccatcat cagcggtcgt    1140
tccggtcgtc tggcgatgaa gataggatct caggattgga atcctggtgc cggctgggat    1200
ctcgtcacct ggggcaacca gtacgcggtt tggagtcaaa aataa                    1245
```

<210> 401
<211> 414
<212> PRT
<213> Unknown

<220>

<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(19)

<400> 401

```
Met Lys Asn Phe Ile Leu Gly Met Ala Ala Met Leu Met Ala Ala Gln
 1               5                  10                  15
Ser Leu Ala Ala Thr Glu Gln Asn Lys Ala Ser Thr Ala Ile Leu Leu
            20                  25                  30
Gln Gly Phe His Trp Asn Ser Ala Asn Gly Ser Trp Tyr Leu Asn Leu
        35                  40                  45
Glu Gly Lys Ala Ala Asp Phe Arg Asp Leu Gly Ile Thr His Val Trp
    50                  55                  60
Phe Pro Pro Pro Ser Asp Ser Gly Ala Arg Glu Gly Tyr Leu Pro Arg
65                  70                  75                  80
Glu Leu Tyr Lys Leu Gln Ser Ala Tyr Gly Ser Glu Glu Gln Leu Arg
                85                  90                  95
Asn Ala Ile Ser Ala Leu Asn Arg Glu Gly Ile Gln Ala Val Ala Asp
            100                 105                 110
Ile Val Ile Asn His Arg Val Gly Ser Trp Asn Trp Gly Asp Phe Asn
        115                 120                 125
Tyr Pro Lys Trp Asp Cys Asp Ser Val Ala Asn Asn Asp Glu Trp Pro
    130                 135                 140
Gly Arg Cys Gly Gly Tyr Asp Ser Gly Ala Ser Tyr Gly Ala Ala Arg
145                 150                 155                 160
Asp Ile Asp His Ser Asn Ala Gln Val Gln Asn Asp Ile Lys Tyr Phe
                165                 170                 175
Leu Ser Glu Arg Leu Arg Gly Val Gly Phe Ser Gly Trp Arg Tyr Asp
            180                 185                 190
Tyr Ser Lys Gly Tyr Ala Pro Tyr Tyr Ala Arg Ile Tyr Asn Glu Ala
        195                 200                 205
Ser Arg Pro Asp Phe Cys Val Gly Glu Val Trp Pro Gly Leu Asn Tyr
    210                 215                 220
Asp Asp Val Asn Ala His Arg Gln Glu Leu Met Asn Tyr Val Asp Gly
225                 230                 235                 240
Thr Gly Gly Ser Cys Gly Ala Phe Asp Phe Thr Ser Lys Gly Leu Leu
                245                 250                 255
Asn Lys Val Leu Arg Asp Asn Asp Tyr Trp Arg Leu Arg Asp Ala Asn
            260                 265                 270
Gly Arg Pro Ala Gly Gly Ile Gly Trp Trp Pro Gln Lys Met Val Thr
        275                 280                 285
Phe Val Asp Asn His Asp Thr Gly Pro Ser Glu Ser Cys Ser Ala Gly
    290                 295                 300
Gln Asn Met Trp Pro Val Pro Cys Asp Lys Val Met Leu Gly Tyr Ala
305                 310                 315                 320
Tyr Ile Leu Thr His Pro Gly Val Pro Ser Leu Tyr Tyr Pro His Leu
                325                 330                 335
Tyr Asp Trp Gly Leu Lys Asn Ser Ile Lys Ser Leu Leu Asn Ala Arg
            340                 345                 350
Arg Ser Val Gly Ile Gln Ser Asn Ser Asn Ile Ser Ile Gln Lys Ala
        355                 360                 365
Glu Asn Gly Leu Tyr Ala Ala Ile Ile Ser Gly Arg Ser Gly Arg Leu
    370                 375                 380
Ala Met Lys Ile Gly Ser Gln Asp Trp Asn Pro Gly Ala Gly Trp Asp
385                 390                 395                 400
Leu Val Thr Trp Gly Asn Gln Tyr Ala Val Trp Ser Gln Lys
                405                 410
```

<210> 402
<211> 1995
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 402

```
atgaactaca ccgcacttgc ggccagccta ctgctgactg cctgcccggc tttggccgcg      60
caaaaaacta ctttcgtgca tttattcgag tggcgctggt ctgatattgc caccgaatgc     120
gaaacatttt taggcccgca gggctttgct gctatccagg tgtcgcctcc gacagagcat     180
attcagggct cggcctggtg gacccgctat caaccggtca gctatgtggt acaaagccgc     240
agtggtaaca gagccgagtt tatctctatg gtgcaacgtt gtaaggcggc aggcgtagat     300
atttatgtcg acgctgtcat caaccatatg gccagcggca gtggtaccgg cgttgctggc     360
agcagttatg gtaacaggca gtaccccatc tacagcagcc ctgattttca taactcctgc     420
accattaaca gcagcgacta tggcaacgac cgctggcggg tgcaaaactg cgaactggtg     480
ggtttacccg atttaaatac cagcgccagc tatgtacaaa atagcctcgc cgcttatctg     540
aacgatatgc tggcgatagg ggtaactggt tttcgtattg acgccgccaa gcatatggcc     600
agttcagata ttgcaggcat caaagccaaa ctcagcagag caccactgat ttaccaggag     660
gtgatcgatc agggcggcga ggtgatcacc gccagcgagt acacaggcaa tggcatggtg     720
actgaattta agtattcggt gcagctgggt aatgtgttta aaaccggccg gctggccagc     780
ctgtctaact ttggtgaagc ctggggtttt ttacctggca atcaggcggt ggtttttgta     840
gataaccacg acaatcagcg cggtcatggc ggtgcaggta atgtcgtgac ctacaaagat     900
ggcaaactct atgatctggc caatgtgttt atgctggcct acccctacgg ttatacccag     960
gtcatgtcca gttacgactt cacggatgca gacgctggcc cgccttcagc tgcagtgcat    1020
aacaatggca gcctgaattg ctttggcagt cagtggaaat gtgaacatcg ctggagcacc    1080
atagcgggcg cagtgcagtt ccgcaacaat acggcggagg aatggcgggt gacccattgg    1140
tgggataacg gcaataacca gatcgccttc ggccgtgctg ctttaggttt tgtcgctatc    1200
aataaagaaa gttacaacct gagcacatca ttacaaaccg gtatgacagc aggcagttac    1260
tgcaatgtgc tcaaaggcag catcagcacc gataaaacca gttgttcagg cgaagtgatc    1320
cagataggca gcgatggccg tattcaggcc aatctggcat cctgggatgc ttttgccatt    1380
catcatcagg cgaaaattgg cggcacagtg cagcctcctg caggcaactg gcagcgtact    1440
gtggtgttta ttcaggccca aactcaaact ggccaggata tgtttgtccg cggtgggctg    1500
gatcatgcag ttgcgcaaag ccagctcggc ttaagctgca gcagtagcaa ttatttgtgc    1560
gccattccta tacggcatcg taatctgaac aacgccacta cagcaccatg gaaagctaac    1620
gatggcttgc tcgactggta tggccgcgaa gccagccaaa gcagcgccgc tgaaggcaca    1680
gcactggact ggaccaccaa cagctggcca gcaagctggg gcacagtccg cactgtggcc    1740
acagatggct atggtgtgga ggtactgaac agctttggcc aacattactg gatgctggat    1800
gtcgatatgg attgcagcaa aaccatcaat ggctggtttg aactcaaagc ttatgtcaaa    1860
aacggccagg gctgggaagg taatatcagc cagagcggca cgccttacgc cagtcagaac    1920
catatggcgc aatgcggtca gatgaatgtg tttagctttg gtcaaaacag cgcacaaatc    1980
agcgccttgc cctag                                                     1995

<210> 403
<211> 664
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(19)

<400> 403
Met Asn Tyr Thr Ala Leu Ala Ala Ser Leu Leu Leu Thr Ala Cys Pro
 1               5                  10                  15
Ala Leu Ala Ala Gln Lys Thr Thr Phe Val His Leu Phe Glu Trp Arg
            20                  25                  30
Trp Ser Asp Ile Ala Thr Glu Cys Glu Thr Phe Leu Gly Pro Gln Gly
        35                  40                  45
Phe Ala Ala Ile Gln Val Ser Pro Pro Thr Glu His Ile Gln Gly Ser
    50                  55                  60
Ala Trp Trp Thr Arg Tyr Gln Pro Val Ser Tyr Val Val Gln Ser Arg
65                  70                  75                  80
Ser Gly Asn Arg Ala Glu Phe Ile Ser Met Val Gln Arg Cys Lys Ala
                85                  90                  95
Ala Gly Val Asp Ile Tyr Val Asp Ala Val Ile Asn His Met Ala Ser
            100                 105                 110
Gly Ser Gly Thr Gly Val Ala Gly Ser Ser Tyr Gly Asn Arg Gln Tyr
        115                 120                 125
Pro Ile Tyr Ser Ser Pro Asp Phe His Asn Ser Cys Thr Ile Asn Ser
    130                 135                 140
Ser Asp Tyr Gly Asn Asp Arg Trp Arg Val Gln Asn Cys Glu Leu Val
145                 150                 155                 160
Gly Leu Pro Asp Leu Asn Thr Ser Ala Ser Tyr Val Gln Asn Ser Leu
```

```
                165                 170                 175
Ala Ala Tyr Leu Asn Asp Met Leu Ala Ile Gly Val Thr Gly Phe Arg
            180                 185                 190
Ile Asp Ala Ala Lys His Met Ala Ser Ser Asp Ile Ala Gly Ile Lys
        195                 200                 205
Ala Lys Leu Ser Arg Ala Pro Leu Ile Tyr Gln Glu Val Ile Asp Gln
    210                 215                 220
Gly Gly Glu Val Ile Thr Ala Ser Glu Tyr Thr Gly Asn Gly Met Val
225                 230                 235                 240
Thr Glu Phe Lys Tyr Ser Val Gln Leu Gly Asn Val Phe Lys Thr Gly
                245                 250                 255
Arg Leu Ala Ser Leu Ser Asn Phe Gly Glu Ala Trp Gly Phe Leu Pro
            260                 265                 270
Gly Asn Gln Ala Val Val Phe Val Asp Asn His Asp Asn Gln Arg Gly
        275                 280                 285
His Gly Gly Ala Gly Asn Val Val Thr Tyr Lys Asp Gly Lys Leu Tyr
    290                 295                 300
Asp Leu Ala Asn Val Phe Met Leu Ala Tyr Pro Tyr Gly Tyr Thr Gln
305                 310                 315                 320
Val Met Ser Ser Tyr Asp Phe Thr Asp Ala Asp Ala Gly Pro Pro Ser
                325                 330                 335
Ala Ala Val His Asn Asn Gly Ser Leu Asn Cys Phe Gly Ser Gln Trp
            340                 345                 350
Lys Cys Glu His Arg Trp Ser Thr Ile Ala Gly Ala Val Gln Phe Arg
        355                 360                 365
Asn Asn Thr Ala Glu Glu Trp Arg Val Thr His Trp Trp Asp Asn Gly
    370                 375                 380
Asn Asn Gln Ile Ala Phe Gly Arg Ala Ala Leu Gly Phe Val Ala Ile
385                 390                 395                 400
Asn Lys Glu Ser Tyr Asn Leu Ser Thr Ser Leu Gln Thr Gly Met Thr
                405                 410                 415
Ala Gly Ser Tyr Cys Asn Val Leu Lys Gly Ser Ile Ser Thr Asp Lys
            420                 425                 430
Thr Ser Cys Ser Gly Glu Val Ile Gln Ile Gly Ser Asp Gly Arg Ile
        435                 440                 445
Gln Ala Asn Leu Ala Ser Trp Asp Ala Phe Ala Ile His His Gln Ala
    450                 455                 460
Lys Ile Gly Gly Thr Val Gln Pro Pro Ala Gly Asn Trp Gln Arg Thr
465                 470                 475                 480
Val Val Phe Ile Gln Ala Gln Thr Gln Thr Gly Gln Asp Met Phe Val
                485                 490                 495
Arg Gly Gly Leu Asp His Ala Val Ala Gln Ser Gln Leu Gly Leu Ser
            500                 505                 510
Cys Ser Ser Ser Asn Tyr Leu Cys Ala Ile Pro Ile Arg His Arg Asn
        515                 520                 525
Leu Asn Asn Ala Thr Thr Ala Pro Trp Lys Ala Asn Asp Gly Leu Leu
    530                 535                 540
Asp Trp Tyr Gly Arg Glu Ala Ser Gln Ser Ser Ala Ala Glu Gly Thr
545                 550                 555                 560
Ala Leu Asp Trp Thr Thr Asn Ser Trp Pro Ala Ser Trp Gly Thr Val
                565                 570                 575
Arg Thr Val Ala Thr Asp Gly Tyr Gly Val Glu Val Leu Asn Ser Phe
            580                 585                 590
Gly Gln His Tyr Trp Met Leu Asp Val Asp Met Asp Cys Ser Lys Thr
        595                 600                 605
Ile Asn Gly Trp Phe Glu Leu Lys Ala Tyr Val Lys Asn Gly Gln Gly
    610                 615                 620
Trp Glu Gly Asn Ile Ser Gln Ser Gly Thr Pro Tyr Ala Ser Gln Asn
625                 630                 635                 640
His Met Ala Gln Cys Gly Gln Met Asn Val Phe Ser Phe Gly Gln Asn
                645                 650                 655
Ser Ala Gln Ile Ser Ala Leu Pro
            660

<210> 404
<211> 1653
<212> DNA
<213> Unknown
```

<220>
<223> Obtained from an environmental sample

<400> 404
atgcgaaaaa tagtccagct cctgatgatc tctgccctac ttgccctaag ctgtgcccct 60
aacaaggagg ccactgccgt tgaaacccct tttgtttggg agggtgccaa tctgtatttc 120
ctcttggtgg accgctttca taacggcaac cccgataacg acaacatcat cccccgaaac 180
gaagaaaccg gggtgctgag aaattttatg ggaggtgacc tgcggggcat cattcaaaaa 240
attgatgaag gctatttttc cgaactaggc atcaacgcca tctggatgac tcccatcgtt 300
gaacaaatcc acggagcggt agatgaaggc accggaaaca cctacggctt ccacggctat 360
tggaccaaag actggaccaa tatcgacccc agctatggca cccgcgaaga cctgaaagaa 420
cttgttgaaa aagcccatgc acaaggcatt cgcatcgtgt tagacgccgt gattaaccac 480
accggaccgg taaccgaact cgatcctgtt tggccggaca cctgggtgcg aaccggcccc 540
aaatgtacct atgacagcta cagcaattat atcgattgta cccttgtcga aaacctgccc 600
gacgtgctta ccgagagcaa agaacccgta gaactccccg aacacctcat tgaaaaatgg 660
aaaaaagaag gccgttacga acaggaaatc gctgaactcg atgccttttt tgcccgtacc 720
ggttatccga gagcaccaaa atattacatc atgaaatggc tttcagacta catcaccgac 780
tttggtattg acggttaccg tgtggacact gtcaaacaca ccaccgaaga tgtgtgggcc 840
gatttctcca aagtatgtaa agacgctttc gcacaataca aaaaagacaa tcccgataag 900
gtacttgacg acaacgactt ctttatgatt ggcgaagtct atggctatgg catccacggg 960
aaaagaatct atgactttgg cgataaacaa gtggattact ttgcccatgg ctttgacaac 1020
ctcatcaact ttgatttcaa aggagatgcc aacaagccgt ttgaagaatt attctctaaa 1080
tacaaccagg ccgtgcaaaa cgacctcaaa ggcgtgagca tcatgaacta cgtttcctca 1140
cacgacgaca gctggccttt tgacaaaaaa agagaacgca gctttgaagc cggcaccaaa 1200
ctattgctta ctcccggcat atcacaggtc tattatggcg atgaaaccgc acgcagtctt 1260
gaaattcccg gcactgtggg cgatgccacc ctgcgttcca tgatgaattg ggacgacctc 1320
caaaacaacc ccaacaccca acaagtgctc aaacacttcc ggctgctcgg gcaattcaga 1380
aaaaaccacc ccgctgtcgg agccggagta catcaggaaa tttccaaagc accgttctat 1440
accttcagca gaacctttac aaaaggagat tacacagata cagtcatcat cggtttagac 1500
ttgccatttg gcaaaaaaga aatacaagtg tcctccattt ttgaagacgg aacagaactg 1560
cacgatgcct tttcaggtat tacaactacc gtaaaggacg gtaaagcagt catcgatagt 1620
caaatgggca ttgtgttatt ggagagaaaa tag 1653

<210> 405
<211> 550
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(26)

<400> 405
Met Arg Lys Ile Val Gln Leu Leu Met Ile Ser Ala Leu Leu Ala Leu
 1   5   10   15
Ser Cys Ala Pro Asn Lys Glu Ala Thr Ala Val Glu Thr Pro Phe Val
            20   25   30
Trp Glu Gly Ala Asn Leu Tyr Phe Leu Leu Val Asp Arg Phe His Asn
        35   40   45
Gly Asn Pro Asp Asn Asp Asn Ile Ile Pro Arg Asn Glu Glu Thr Gly
    50   55   60
Val Leu Arg Asn Phe Met Gly Gly Asp Leu Arg Gly Ile Ile Gln Lys
65   70   75   80
Ile Asp Glu Gly Tyr Phe Ser Glu Leu Gly Ile Asn Ala Ile Trp Met
                85   90   95
Thr Pro Ile Val Glu Gln Ile His Gly Ala Val Asp Glu Gly Thr Gly
            100   105   110
Asn Thr Tyr Gly Phe His Gly Tyr Trp Thr Lys Asp Trp Thr Asn Ile
        115   120   125
Asp Pro Ser Tyr Gly Thr Arg Glu Asp Leu Lys Glu Leu Val Glu Lys
    130   135   140
Ala His Ala Gln Gly Ile Arg Ile Val Leu Asp Ala Val Ile Asn His
145   150   155   160
Thr Gly Pro Val Thr Glu Leu Asp Pro Val Trp Pro Asp Thr Trp Val
                165   170   175

```
Arg Thr Gly Pro Lys Cys Thr Tyr Asp Ser Tyr Ser Asn Tyr Ile Asp
            180                 185                 190
Cys Thr Leu Val Glu Asn Leu Pro Asp Val Leu Thr Glu Ser Lys Glu
        195                 200                 205
Pro Val Glu Leu Pro Glu His Leu Ile Glu Lys Trp Lys Lys Glu Gly
    210                 215                 220
Arg Tyr Glu Gln Glu Ile Ala Glu Leu Asp Ala Phe Phe Ala Arg Thr
225                 230                 235                 240
Gly Tyr Pro Arg Ala Pro Lys Tyr Tyr Ile Met Lys Trp Leu Ser Asp
                245                 250                 255
Tyr Ile Thr Asp Phe Gly Ile Asp Gly Tyr Arg Val Asp Thr Val Lys
            260                 265                 270
His Thr Thr Glu Asp Val Trp Ala Asp Phe Ser Lys Val Cys Lys Asp
        275                 280                 285
Ala Phe Ala Gln Tyr Lys Lys Asp Asn Pro Asp Lys Val Leu Asp Asp
    290                 295                 300
Asn Asp Phe Phe Met Ile Gly Glu Val Tyr Gly Tyr Gly Ile His Gly
305                 310                 315                 320
Lys Arg Ile Tyr Asp Phe Gly Asp Lys Gln Val Asp Tyr Phe Ala His
                325                 330                 335
Gly Phe Asp Asn Leu Ile Asn Phe Asp Phe Lys Gly Asp Ala Asn Lys
            340                 345                 350
Pro Phe Glu Glu Leu Phe Ser Lys Tyr Asn Gln Ala Val Gln Asn Asp
        355                 360                 365
Leu Lys Gly Val Ser Ile Met Asn Tyr Val Ser Ser His Asp Asp Ser
    370                 375                 380
Trp Pro Phe Asp Lys Lys Arg Glu Arg Ser Phe Glu Ala Gly Thr Lys
385                 390                 395                 400
Leu Leu Leu Thr Pro Gly Ile Ser Gln Val Tyr Tyr Gly Asp Glu Thr
                405                 410                 415
Ala Arg Ser Leu Glu Ile Pro Gly Thr Val Gly Asp Ala Thr Leu Arg
            420                 425                 430
Ser Met Met Asn Trp Asp Asp Leu Gln Asn Asn Pro Asn Thr Gln Gln
        435                 440                 445
Val Leu Lys His Phe Arg Leu Leu Gly Gln Phe Arg Lys Asn His Pro
    450                 455                 460
Ala Val Gly Ala Gly Val His Gln Glu Ile Ser Lys Ala Pro Phe Tyr
465                 470                 475                 480
Thr Phe Ser Arg Thr Phe Thr Lys Gly Asp Tyr Thr Asp Thr Val Ile
                485                 490                 495
Ile Gly Leu Asp Leu Pro Phe Gly Lys Lys Glu Ile Gln Val Ser Ser
            500                 505                 510
Ile Phe Glu Asp Gly Thr Glu Leu His Asp Ala Phe Ser Gly Ile Thr
        515                 520                 525
Thr Thr Val Lys Asp Gly Lys Ala Val Ile Asp Ser Gln Met Gly Ile
    530                 535                 540
Val Leu Leu Glu Arg Lys
545                 550
```

<210> 406
<211> 1398
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 406

```
atgaaaaaca cagcgggaat ctgggtgatc gcaggcatgc tgatcgcccc actggcgcaa      60
gccgatgtca tactgcacgc cttcaactgg aaatacagcg aggtcaccgc caaggccgac     120
ctcatcaagg gggcgggcta caagcaggtg ctcatctcgc cccctctcaa gtcctccggc     180
aacgagtggt gggcccgtta ccagccccag gatctgcggc tcgtcgacag cccgctgggc     240
aacaagcagg atctcgagca gctcatcgcc gccatgcagg cccgcggcat cgccgtctat     300
gcggatgtgg tgctcaacca catggccaac gagagctgga agcgcagtga cctcaactac     360
cccggcagcg agctgttgca aacctatgcc gccaaccccg cctacttcga gcggcagaag     420
ctgttcggcg atctggcgca aaaccttctg tcgggtcagg acttccaccc cgagggctgc     480
atcaccgact ggagcgatcc cggtcatgtc cagtactggc gcctgtgtgg cggcgccggc     540
gacaaggggt tgccggatct cgaccccaac aactgggtgg tgagccagca gcaggcctat     600
```

ctcaaggcgc tcaaggggat ggggatcaag ggctttcggg tcgatgcggt caagcacatg 660
agcgattacc agatcaacgc cgtcttcacc cccgagatca agcagggcat gcacgtgttt 720
ggggaggtga tcacctccgg gggcgccggc agcacggatt acgagcgctt cctcaagccc 780
tatctcgaca acagcggcca gggggcctat gacttcccgc tgtttgcctc cctgcgcggg 840
gcgctcggct atggcggcag catgaatcag ctggcggatc ccggcgccta cggccaggcc 900
ctgccgggca accgggcggt caccttcgcc atcacccacg acatccccac caacgacggc 960
ttccgctacc agatcctgaa ccagaccgac gagaagctgg cctatgccta tctgctcggc 1020
cgcgatgggg gctcccccct ggtctattcg gatcacggcg agacccagga caaggacggc 1080
ctgcgctggc aggactacta cctgcgcagc gatctcaagg ggatgatccg cttccacaac 1140
gccgtgcagg gtcagcccat gcagctcatc ggcagcggcg actgcttcgt gctgttcaag 1200
cgcggcaagc aggggctggt cggggtcaac aagtgtgact acgagcagga gtactggctc 1260
gacaccgcca ggttcgagct gaactggtat cgcaactaca gggacgtgct ggatcaaaat 1320
gccgtcatca atgtgcagag ccagtgggta cgggtcgcca taccggcgcg cagcgcccgc 1380
ctctggctgc aggaataa 1398

<210> 407
<211> 465
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(21)

<400> 407
Met Lys Asn Thr Ala Gly Ile Trp Val Ile Ala Gly Met Leu Ile Ala
 1   5   10   15
Pro Leu Ala Gln Ala Asp Val Ile Leu His Ala Phe Asn Trp Lys Tyr
            20   25   30
Ser Glu Val Thr Ala Lys Ala Asp Leu Ile Lys Gly Ala Gly Tyr Lys
        35   40   45
Gln Val Leu Ile Ser Pro Pro Leu Lys Ser Ser Gly Asn Glu Trp Trp
    50   55   60
Ala Arg Tyr Gln Pro Gln Asp Leu Arg Leu Val Asp Ser Pro Leu Gly
65   70   75   80
Asn Lys Gln Asp Leu Glu Gln Leu Ile Ala Ala Met Gln Ala Arg Gly
                85   90   95
Ile Ala Val Tyr Ala Asp Val Val Leu Asn His Met Ala Asn Glu Ser
            100   105   110
Trp Lys Arg Ser Asp Leu Asn Tyr Pro Gly Ser Glu Leu Leu Gln Thr
        115   120   125
Tyr Ala Ala Asn Pro Ala Tyr Phe Glu Arg Gln Lys Leu Phe Gly Asp
    130   135   140
Leu Ala Gln Asn Leu Leu Ser Gly Gln Asp Phe His Pro Glu Gly Cys
145   150   155   160
Ile Thr Asp Trp Ser Asp Pro Gly His Val Gln Tyr Trp Arg Leu Cys
                165   170   175
Gly Gly Ala Gly Asp Lys Gly Leu Pro Asp Leu Asp Pro Asn Asn Trp
            180   185   190
Val Val Ser Gln Gln Gln Ala Tyr Leu Lys Ala Leu Lys Gly Met Gly
        195   200   205
Ile Lys Gly Phe Arg Val Asp Ala Val Lys His Met Ser Asp Tyr Gln
    210   215   220
Ile Asn Ala Val Phe Thr Pro Glu Ile Lys Gln Gly Met His Val Phe
225   230   235   240
Gly Glu Val Ile Thr Ser Gly Gly Ala Gly Ser Thr Asp Tyr Glu Arg
                245   250   255
Phe Leu Lys Pro Tyr Leu Asp Asn Ser Gly Gln Gly Ala Tyr Asp Phe
            260   265   270
Pro Leu Phe Ala Ser Leu Arg Gly Ala Leu Gly Tyr Gly Gly Ser Met
        275   280   285
Asn Gln Leu Ala Asp Pro Gly Ala Tyr Gly Gln Ala Leu Pro Gly Asn
    290   295   300
Arg Ala Val Thr Phe Ala Ile Thr His Asp Ile Pro Thr Asn Asp Gly
305   310   315   320
Phe Arg Tyr Gln Ile Leu Asn Gln Thr Asp Glu Lys Leu Ala Tyr Ala

325 330 335
Tyr Leu Leu Gly Arg Asp Gly Gly Ser Pro Leu Val Tyr Ser Asp His
340 345 350
Gly Glu Thr Gln Asp Lys Asp Gly Leu Arg Trp Gln Asp Tyr Tyr Leu
355 360 365
Arg Ser Asp Leu Lys Gly Met Ile Arg Phe His Asn Ala Val Gln Gly
370 375 380
Gln Pro Met Gln Leu Ile Gly Ser Gly Asp Cys Phe Val Leu Phe Lys
385 390 395 400
Arg Gly Lys Gln Gly Leu Val Gly Val Asn Lys Cys Asp Tyr Glu Gln
405 410 415
Glu Tyr Trp Leu Asp Thr Ala Arg Phe Glu Leu Asn Trp Tyr Arg Asn
420 425 430
Tyr Arg Asp Val Leu Asp Gln Asn Ala Val Ile Asn Val Gln Ser Gln
435 440 445
Trp Val Arg Val Ala Ile Pro Ala Arg Ser Ala Arg Leu Trp Leu Gln
450 455 460
Glu
465

<210> 408
<211> 1476
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 408
atgccgtcga tgcccccgtt gcacgcttca gcgtcagccc cgctggcccc gtcccccgct 60
gcaccgcgcg cgacgtcgcg cctgcgccgt ggcctggtac gcctcgccgc caccgcactg 120
ctggccttcg gtgccatcgg cggcgcgcag gccgatgtga tcctgcatgc cttcaactgg 180
cgctatgccg acgtggaggc ccgcgcctcg cagatccagg ccgcgggcta ccgcgcggtg 240
ctggtggccc cggcgtacaa gagcgagggc agcgcctggt gggcgcgcta ccagccgcag 300
gattaccggc tgatccatca cccgctgggc gacaccaccg cgttccgcag catggtcaat 360
gcgctgaatg cgcgcggcat ccgcgtgtac gcggacgtca tcctaaacca catggccaac 420
gagtcggcgc aacgcagcga cctgaactac ccgggccagc gcctgctcaa cctctacgcc 480
ggcaacacca gctactacaa cggccagcgg ctgttcggcg acctgcgcta caacttcctc 540
agcagctggg atttcggcca ggccacgtgc atcagcgact acaacgacgt gtggcaggtg 600
cagaactggc gcctgtgcgg cggcggtggt gacgccggcc tgccggacct cgtcaacggc 660
gactacgtgg tgtcgcagca gcaggcctac ctgcgcgcgc tgaaggccat gggcgtgaag 720
ggcttccgca tcgacgccgc gaagcacatg cccatcgccc acctcaaccg ggtgctgacc 780
agcgacatca agtcgggcat gcacatcttc ggcgagatca tcactggcgg cggcgcgggc 840
aacggtgagt acgaccgttt cctcgcgccg tacctgcagt acaccgacca cggcgcgtac 900
gacttcccgc tgcacacgca gatccggggc gcattcggct tcggtggcag catgagcaac 960
ctggtcgacc cgcaggccta cgggcaggcg ctggcggcgg cgcgcgcggt caccttcacc 1020
gtcacgcacg acatcccgaa caacagcggc ttccgctacg cgctgatgga ccccaccgac 1080
gaaaccctcg cgtacgccta cgtgatgggg cgcaacggcg gctccccgct gctgtactcg 1140
gacaacaacg agagcggcga caaccgttgg gtcaatgcgt ggaaccgtac cgacctcaag 1200
cagatggtgc gcttccacaa tgcggtgcag ggcagcgacc aggcggtgct ggcgcacggc 1260
agctgctacc tgctgttccg ccgcggcagc ctgggcatcg tcggcatcaa caagtgcggc 1320
agcccggtca acgtcaccgt caacatgaac aacagcgtgc tgtggtggta cgccaactac 1380
accgacgtgc tcgattcgac cagcaaggtc tatatcagca gcggaagcca taccttcacc 1440
ctgccggcgc gccgcgcgcg cctgtggctg aggtaa 1476

<210> 409
<211> 491
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(51)

<400> 409
Met Pro Ser Met Pro Pro Leu His Ala Ser Ala Ser Ala Pro Leu Ala

```
 1               5                  10                  15
Pro Ser Pro Ala Ala Pro Arg Ala Thr Ser Arg Leu Arg Arg Gly Leu
            20                  25                  30
Val Arg Leu Ala Ala Thr Ala Leu Leu Ala Phe Gly Ala Ile Gly Gly
        35                  40                  45
Ala Gln Ala Asp Val Ile Leu His Ala Phe Asn Trp Arg Tyr Ala Asp
    50                  55                  60
Val Glu Ala Arg Ala Ser Gln Ile Gln Ala Ala Gly Tyr Arg Ala Val
65                  70                  75                  80
Leu Val Ala Pro Ala Tyr Lys Ser Glu Gly Ser Ala Trp Trp Ala Arg
                85                  90                  95
Tyr Gln Pro Gln Asp Tyr Arg Leu Ile His His Pro Leu Gly Asp Thr
            100                 105                 110
Thr Ala Phe Arg Ser Met Val Asn Ala Leu Asn Ala Arg Gly Ile Arg
        115                 120                 125
Val Tyr Ala Asp Val Ile Leu Asn His Met Ala Asn Glu Ser Ala Gln
    130                 135                 140
Arg Ser Asp Leu Asn Tyr Pro Gly Gln Arg Leu Leu Asn Leu Tyr Ala
145                 150                 155                 160
Gly Asn Thr Ser Tyr Tyr Asn Gly Gln Arg Leu Phe Gly Asp Leu Arg
                165                 170                 175
Tyr Asn Phe Leu Ser Ser Trp Asp Phe Gly Gln Ala Thr Cys Ile Ser
            180                 185                 190
Asp Tyr Asn Asp Val Trp Gln Val Gln Asn Trp Arg Leu Cys Gly Gly
        195                 200                 205
Gly Gly Asp Ala Gly Leu Pro Asp Leu Val Asn Gly Asp Tyr Val Val
    210                 215                 220
Ser Gln Gln Gln Ala Tyr Leu Arg Ala Leu Lys Ala Met Gly Val Lys
225                 230                 235                 240
Gly Phe Arg Ile Asp Ala Ala Lys His Met Pro Ile Ala His Leu Asn
                245                 250                 255
Arg Val Leu Thr Ser Asp Ile Lys Ser Gly Met His Ile Phe Gly Glu
            260                 265                 270
Ile Ile Thr Gly Gly Gly Ala Gly Asn Gly Glu Tyr Asp Arg Phe Leu
        275                 280                 285
Ala Pro Tyr Leu Gln Tyr Thr Asp His Gly Ala Tyr Asp Phe Pro Leu
    290                 295                 300
His Thr Gln Ile Arg Gly Ala Phe Gly Phe Gly Gly Ser Met Ser Asn
305                 310                 315                 320
Leu Val Asp Pro Gln Ala Tyr Gly Gln Ala Leu Ala Ala Ala Arg Ala
                325                 330                 335
Val Thr Phe Thr Val Thr His Asp Ile Pro Asn Asn Ser Gly Phe Arg
            340                 345                 350
Tyr Ala Leu Met Asp Pro Thr Asp Glu Thr Leu Ala Tyr Ala Tyr Val
        355                 360                 365
Met Gly Arg Asn Gly Gly Ser Pro Leu Leu Tyr Ser Asp Asn Asn Glu
    370                 375                 380
Ser Gly Asp Asn Arg Trp Val Asn Ala Trp Asn Arg Thr Asp Leu Lys
385                 390                 395                 400
Gln Met Val Arg Phe His Asn Ala Val Gln Gly Ser Asp Gln Ala Val
                405                 410                 415
Leu Ala His Gly Ser Cys Tyr Leu Leu Phe Arg Arg Gly Ser Leu Gly
            420                 425                 430
Ile Val Gly Ile Asn Lys Cys Gly Ser Pro Val Asn Val Thr Val Asn
        435                 440                 445
Met Asn Asn Ser Val Leu Trp Trp Tyr Ala Asn Tyr Thr Asp Val Leu
    450                 455                 460
Asp Ser Thr Ser Lys Val Tyr Ile Ser Ser Gly Ser His Thr Phe Thr
465                 470                 475                 480
Leu Pro Ala Arg Arg Ala Arg Leu Trp Leu Arg
                485                 490
```

<210> 410
<211> 1875
<212> DNA
<213> Unknown

<220>

<223> Obtained from an environmental sample

<400> 410
atgaaaaata caagcctgac cctgttagca ggctgcatgg ccctggcgtt taacgcccag 60
gccgcagtca gccaaaataa ccctgacatc atgttgcaag gcttccactg gaactccgcc 120
aaagcgggtg gctggtacaa caccctgcaa ggcaatgtga acgaaatcgc cagcgccggc 180
ttcaacatgg tgtggctgcc gccccctcg caggccggct cactggaagg ttacctgccg 240
gagcagtaca acaacctcaa ctccaactac ggcaccgaaa ctcagctgcg caacctgctc 300
agcgccctca aggccaacaa cgtcaaggcc attgccgaca tcgtcatcaa ccaccgtaat 360
ggctctggca gctggtgtac cttcaccaac ccggcctggg gttttgatgc catcgtctct 420
aacgacgaag cctggggcgc agccggctcc aactgcagcg gcacccgtgg cgctgccgac 480
tctggtgacg gctatcacgc cgcccgcgac atcgaccaca gcaaaaccta tgtgcgcgac 540
agccttaaag agtggatgaa cgtccgcctc aaaggcgttg gctttgacgg ctggcgctat 600
gactacgtca aagggttcag cggtgtctac gtcggcgaat ataacgccgc caccaaccct 660
tacttcagcg tcggtgaata ctggacctcg ctctgctaca acggcgaaga gtgctttcct 720
ggtggctctt ctccagaatc tcaccgccag gcccagatca actgggtcga caagaccaat 780
ggcaacagcg cggtctttga cttcaccacc aaagggctgc tcaacaaagc gctggagacc 840
tacaactaca gccacctgcg cgatgccaac ggcaaaccgg ccggggtgat cggctactgg 900
ccgtcgcgcg ccgtgacctt cgtcgacaac cacgacactg gcccgagcga agtctgtggc 960
aatgcccaga acaactggcc ggtgccgtgc gacaaggtga tgcagggtta tgcctacatc 1020
ctcacccacc cgggtgtgcc ctcggtctac tacgcccact acttcaactg ggggctgggc 1080
agcgaaatca agaagttgat gaaactgcgc aaagacatgg ggctgcactc cgactccgcc 1140
gtgaccatcg acaaggcgca gcagggcctc tatgccgctt acatcggtgg caaggtggca 1200
gtcaaactgg gtaatggttc ctggtcgccg tcaggcgctg gctggaccct ggcccagagc 1260
ggcactgatt gggccgtgtg gaaaaaagat gacggcaata acttcaagcg cactgtggtg 1320
ctgatctacg gcgaaactgc tgccggtcag gacatgttca tccgcggtgg tatcgaccac 1380
gcctatgccg cagccaacct cggcaagacc tgcaccagca ctaactacga gtgcgccatt 1440
ccgatcaccc acaacaacct gcgcaacgcc accaccgcac cgtggaaggc taacgacaag 1500
taccttgact ggtacgacgg tgtcgaagcc ggccagagca ctgccgctca gggctctgcc 1560
gctgactgga ccaccaacgt gtggccgtcg acctggggcg ccgtcaagac tgtggctgcc 1620
gatggtttcg gtgtagaacc gctcaacacc tacggtcagc actactggat gctggatgta 1680
cagatggatt gctccaagac cgtccaaggc atctggttcg agttcaagac cttcatcagc 1740
aacggcccgg gttgggaagc caacgtcgcc cagagcggta ccccgtacgc cagcggtaac 1800
cactttggcc agtgcggtaa ggtcaacgtg ttcaagcgcg gcgtcagcgc cccggtggct 1860
atcagagact tctga 1875

<210> 411
<211> 624
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(21)

<400> 411
Met Lys Asn Thr Ser Leu Thr Leu Leu Ala Gly Cys Met Ala Leu Ala
 1   5   10   15
Phe Asn Ala Gln Ala Ala Val Ser Gln Asn Asn Pro Asp Ile Met Leu
            20   25   30
Gln Gly Phe His Trp Asn Ser Ala Lys Ala Gly Gly Trp Tyr Asn Thr
        35   40   45
Leu Gln Gly Asn Val Asn Glu Ile Ala Ser Ala Gly Phe Asn Met Val
    50   55   60
Trp Leu Pro Pro Pro Ser Gln Ala Gly Ser Leu Glu Gly Tyr Leu Pro
65   70   75   80
Glu Gln Tyr Asn Asn Leu Asn Ser Asn Tyr Gly Thr Glu Thr Gln Leu
                85   90   95
Arg Asn Leu Leu Ser Ala Leu Lys Ala Asn Asn Val Lys Ala Ile Ala
            100   105   110
Asp Ile Val Ile Asn His Arg Asn Gly Ser Gly Ser Trp Cys Thr Phe
        115   120   125
Thr Asn Pro Ala Trp Gly Phe Asp Ala Ile Val Ser Asn Asp Glu Ala
    130   135   140
Trp Gly Ala Ala Gly Ser Asn Cys Ser Gly Thr Arg Gly Ala Ala Asp
145   150   155   160

```
Ser Gly Asp Gly Tyr His Ala Ala Arg Asp Ile Asp His Ser Lys Thr
                165                 170                 175
Tyr Val Arg Asp Ser Leu Lys Glu Trp Met Asn Val Arg Leu Lys Gly
            180                 185                 190
Val Gly Phe Asp Gly Trp Arg Tyr Asp Tyr Val Lys Gly Phe Ser Gly
        195                 200                 205
Val Tyr Val Gly Glu Tyr Asn Ala Ala Thr Asn Pro Tyr Phe Ser Val
    210                 215                 220
Gly Glu Tyr Trp Thr Ser Leu Cys Tyr Asn Gly Glu Glu Cys Phe Pro
225                 230                 235                 240
Gly Gly Ser Ser Pro Glu Ser His Arg Gln Ala Gln Ile Asn Trp Val
                245                 250                 255
Asp Lys Thr Asn Gly Asn Ser Ala Val Phe Asp Phe Thr Thr Lys Gly
            260                 265                 270
Leu Leu Asn Lys Ala Leu Glu Thr Tyr Asn Tyr Ser His Leu Arg Asp
        275                 280                 285
Ala Asn Gly Lys Pro Ala Gly Val Ile Gly Tyr Trp Pro Ser Arg Ala
    290                 295                 300
Val Thr Phe Val Asp Asn His Asp Thr Gly Pro Ser Glu Val Cys Gly
305                 310                 315                 320
Asn Ala Gln Asn Asn Trp Pro Val Pro Cys Asp Lys Val Met Gln Gly
                325                 330                 335
Tyr Ala Tyr Ile Leu Thr His Pro Gly Val Pro Ser Val Tyr Tyr Ala
            340                 345                 350
His Tyr Phe Asn Trp Gly Leu Gly Ser Glu Ile Lys Lys Leu Met Lys
        355                 360                 365
Leu Arg Lys Asp Met Gly Leu His Ser Asp Ser Ala Val Thr Ile Asp
    370                 375                 380
Lys Ala Gln Gln Gly Leu Tyr Ala Ala Tyr Ile Gly Gly Lys Val Ala
385                 390                 395                 400
Val Lys Leu Gly Asn Gly Ser Trp Ser Pro Ser Gly Ala Gly Trp Thr
                405                 410                 415
Leu Ala Gln Ser Gly Thr Asp Trp Ala Val Trp Lys Lys Asp Asp Gly
            420                 425                 430
Asn Asn Phe Lys Arg Thr Val Val Leu Ile Tyr Gly Glu Thr Ala Ala
        435                 440                 445
Gly Gln Asp Met Phe Ile Arg Gly Gly Ile Asp His Ala Tyr Ala Ala
    450                 455                 460
Ala Asn Leu Gly Lys Thr Cys Thr Ser Thr Asn Tyr Glu Cys Ala Ile
465                 470                 475                 480
Pro Ile Thr His Asn Asn Leu Arg Asn Ala Thr Thr Ala Pro Trp Lys
                485                 490                 495
Ala Asn Asp Lys Tyr Leu Asp Trp Tyr Asp Gly Val Glu Ala Gly Gln
            500                 505                 510
Ser Thr Ala Ala Gln Gly Ser Ala Ala Asp Trp Thr Thr Asn Val Trp
        515                 520                 525
Pro Ser Thr Trp Gly Ala Val Lys Thr Val Ala Ala Asp Gly Phe Gly
    530                 535                 540
Val Glu Pro Leu Asn Thr Tyr Gly Gln His Tyr Trp Met Leu Asp Val
545                 550                 555                 560
Gln Met Asp Cys Ser Lys Thr Val Gln Gly Ile Trp Phe Glu Phe Lys
                565                 570                 575
Thr Phe Ile Ser Asn Gly Pro Gly Trp Glu Ala Asn Val Ala Gln Ser
            580                 585                 590
Gly Thr Pro Tyr Ala Ser Gly Asn His Phe Gly Gln Cys Gly Lys Val
        595                 600                 605
Asn Val Phe Lys Arg Gly Val Ser Ala Pro Val Ala Ile Arg Asp Phe
    610                 615                 620

<210> 412
<211> 2088
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 412
```

```
gtggacgcac gcaccggcga gcggcgtgcc gcgttggccg cgctcccggg gcgcgagcgc      60
gtcgtgctcg agcgcgcgac gccggagatc gacggcgggc gcttcgccgc gaagcgcgtg     120
ctcggtgacg ccgtcgtcgt cgaggtcgat gcgttcgcgg acggccacga ccgcgtcgcg     180
tgcgcgctgc ggtggcggca cgaggaagac gccgcgtggc gcgaggcgcc gatggaggcg     240
ctcgcgaacg atcggtggcg tgcgacgttc gcggtcgatc gcctcgggcg ctggcggtat     300
gcgttcgccg cgtggatcga tcgctacggg acgtgggtgc acgacctcgc ccggcgcccg     360
ctcgccgacg ccgatctgcc cgcgacgttc gcagccggag cggcgctcgt gcgcgaggcc     420
gccgcgggcg cgcaagggca cgcggccgag cgactgcgcg aggtcgccgc gctgctcggc     480
ggcgaggccg cggagagcgc gaagcgcgcg gtcgcgtcga gtgcggacac gtgcgcgctg     540
atgcgcgagc acgacccgcg gctgcacgcg gcggactacg gtcgcgtgct ggaggtgatc     600
gtcgatccgc cgctcgcgcg ttgcagtgcg tggtacgagt tcttcccgcg ctcgttcggc     660
gtcgcagagg gagcggcgcc aggcacgccc gcgcgaacgc cgccgagtgc gtcggcgacc     720
ggcgcgttgc gccacggccg gttccgcgac ctcgagccga tgctcgacta cgtcgcgcgg     780
ctcggcttcg acgtgctgca cttcccgccg atccatccga tcggccgcgc gttccgcaag     840
ggtccgaaca acgcgctcgc ggcgggaccc gaggacgtcg gcagcccgtg ggcgatcggg     900
ctgcgcggcg cgggcggcga gggcgggcat cgcgccgtgc acccggagct cggcacgctc     960
gaggacttcc gctggctcgt ggagcgcgcc cgagcgcgcg gcatcgagat cgcgctcgac    1020
gtcgcgttcc agtgctcgcc ggaccacccg tacgtgcacg agcacccgga gtggttccgt    1080
cgccggccgg acggcagcat ccagtacgcc gagaacccgc cgaagaagta ccaggacatc    1140
tacccgttcg acttcgagtc cgagcactgg cccgcgctgt gggacgaact cgcgggcgtg    1200
ttcgagttct gggtcgagca gggcgtgacg atcttccgcg tcgacaaccc gcacacgaag    1260
gcgttcccgt tctgggagta cgcgatcgcg cgcgtgaaga gtcgctgccc gcaggcgatc    1320
ttcctgtccg aggcattcac gcggccgaag gtcatgcacc ggctcgcgaa gctcggcttc    1380
acgcagtcgt acacgtactt cacctggcgc aacacgcggc acgagctcgt gcagtacttc    1440
accgagctcg cgcacgggcc gggagcggag tacttccggc cgaacgcgtg gccgaacacg    1500
ccggacatcc tgccggagta cctgcagctc ggcggccgcg cggcattcat ggtgcgcgcg    1560
gtgctcgcgt cgctgctgtc gtcgaactgg ggcgtgtacg gaccggcatt cgagctgctc    1620
gagcacgagc cgcgttccgc cggcagcgag gagtacctcg actccgagaa gtaccagctg    1680
cggcggtggg atctcgaccg gccggacagc ctcgcggagt tcgtcgccgt gctgaaccgc    1740
atccggcgcg agtcgccggc gctgcagcag atggcgacgc tcgcgttcca cgacgccgac    1800
aacgagcagc tgctgtgctg gtcgaagacc gccggcgacg acgcgatcgt cgtcgtcgtg    1860
aacctcgacc cgcaccacgc gcaggccggc tggatcacgc tgccgctcga gcgactcggg    1920
ctcgagccgg agcgcgcgta ccaggtgcac gacctgctcg gcggcggacg cttcctctgg    1980
agcgggccgc ggaactacgt cgcgctcgcg ccggacgcgc tgccggcgca cgtgttccgc    2040
gtgcggcgac atgcgcgccg cgagcaggac ttcgactact tcctctga                 2088

<210> 413
<211> 695
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 413
Met Asp Ala Arg Thr Gly Glu Arg Arg Ala Ala Leu Ala Ala Leu Pro
 1               5                  10                  15
Gly Arg Glu Arg Val Val Leu Glu Arg Ala Thr Pro Glu Ile Asp Gly
            20                  25                  30
Gly Arg Phe Ala Ala Lys Arg Val Leu Gly Asp Ala Val Val Val Glu
        35                  40                  45
Val Asp Ala Phe Ala Asp Gly His Asp Arg Val Ala Cys Ala Leu Arg
    50                  55                  60
Trp Arg His Glu Glu Asp Ala Ala Trp Arg Glu Ala Pro Met Glu Ala
65                  70                  75                  80
Leu Ala Asn Asp Arg Trp Arg Ala Thr Phe Ala Val Asp Arg Leu Gly
                85                  90                  95
Arg Trp Arg Tyr Ala Phe Ala Ala Trp Ile Asp Arg Tyr Gly Thr Trp
            100                 105                 110
Val His Asp Leu Ala Arg Arg Pro Leu Ala Asp Ala Asp Leu Pro Ala
        115                 120                 125
Thr Phe Ala Ala Gly Ala Ala Leu Val Arg Glu Ala Ala Ala Gly Ala
    130                 135                 140
Gln Gly His Ala Ala Glu Arg Leu Arg Glu Val Ala Ala Leu Leu Gly
145                 150                 155                 160
Gly Glu Ala Ala Glu Ser Ala Lys Arg Ala Val Ala Ser Ser Ala Asp
                165                 170                 175
Thr Cys Ala Leu Met Arg Glu His Asp Pro Arg Leu His Ala Ala Asp
```

```
            180                 185                 190
Tyr Gly Arg Val Leu Glu Val Ile Val Asp Pro Pro Leu Ala Arg Cys
        195                 200                 205
Ser Ala Trp Tyr Glu Phe Phe Pro Arg Ser Phe Gly Val Ala Glu Gly
    210                 215                 220
Ala Ala Pro Gly Thr Pro Ala Arg Thr Pro Pro Ser Ala Ser Ala Thr
225                 230                 235                 240
Gly Ala Leu Arg His Gly Arg Phe Arg Asp Leu Glu Pro Met Leu Asp
                245                 250                 255
Tyr Val Ala Arg Leu Gly Phe Asp Val Leu His Phe Pro Pro Ile His
            260                 265                 270
Pro Ile Gly Arg Ala Phe Arg Lys Gly Pro Asn Asn Ala Leu Ala Ala
        275                 280                 285
Gly Pro Glu Asp Val Gly Ser Pro Trp Ala Ile Gly Leu Arg Gly Ala
    290                 295                 300
Gly Gly Glu Gly Gly His Arg Ala Val His Pro Glu Leu Gly Thr Leu
305                 310                 315                 320
Glu Asp Phe Arg Trp Leu Val Glu Arg Ala Arg Ala Arg Gly Ile Glu
                325                 330                 335
Ile Ala Leu Asp Val Ala Phe Gln Cys Ser Pro Asp His Pro Tyr Val
            340                 345                 350
His Glu His Pro Glu Trp Phe Arg Arg Arg Pro Asp Gly Ser Ile Gln
        355                 360                 365
Tyr Ala Glu Asn Pro Pro Lys Lys Tyr Gln Asp Ile Tyr Pro Phe Asp
    370                 375                 380
Phe Glu Ser Glu His Trp Pro Ala Leu Trp Asp Glu Leu Ala Gly Val
385                 390                 395                 400
Phe Glu Phe Trp Val Glu Gln Gly Val Thr Ile Phe Arg Val Asp Asn
                405                 410                 415
Pro His Thr Lys Ala Phe Pro Phe Trp Glu Tyr Ala Ile Ala Arg Val
            420                 425                 430
Lys Ser Arg Cys Pro Gln Ala Ile Phe Leu Ser Glu Ala Phe Thr Arg
        435                 440                 445
Pro Lys Val Met His Arg Leu Ala Lys Leu Gly Phe Thr Gln Ser Tyr
    450                 455                 460
Thr Tyr Phe Thr Trp Arg Asn Thr Arg His Glu Leu Val Gln Tyr Phe
465                 470                 475                 480
Thr Glu Leu Ala His Gly Pro Gly Ala Glu Tyr Phe Arg Pro Asn Ala
                485                 490                 495
Trp Pro Asn Thr Pro Asp Ile Leu Pro Glu Tyr Leu Gln Leu Gly Gly
            500                 505                 510
Arg Ala Ala Phe Met Val Arg Ala Val Leu Ala Ser Leu Leu Ser Ser
        515                 520                 525
Asn Trp Gly Val Tyr Gly Pro Ala Phe Glu Leu Leu Glu His Glu Pro
    530                 535                 540
Arg Ser Ala Gly Ser Glu Glu Tyr Leu Asp Ser Glu Lys Tyr Gln Leu
545                 550                 555                 560
Arg Arg Trp Asp Leu Asp Arg Pro Asp Ser Leu Ala Glu Phe Val Ala
                565                 570                 575
Val Leu Asn Arg Ile Arg Arg Glu Ser Pro Ala Leu Gln Gln Met Ala
            580                 585                 590
Thr Leu Ala Phe His Asp Ala Asp Asn Glu Gln Leu Leu Cys Trp Ser
        595                 600                 605
Lys Thr Ala Gly Asp Asp Ala Ile Val Val Val Val Asn Leu Asp Pro
    610                 615                 620
His His Ala Gln Ala Gly Trp Ile Thr Leu Pro Leu Glu Arg Leu Gly
625                 630                 635                 640
Leu Glu Pro Glu Arg Ala Tyr Gln Val His Asp Leu Leu Gly Gly Gly
                645                 650                 655
Arg Phe Leu Trp Ser Gly Pro Arg Asn Tyr Val Ala Leu Ala Pro Asp
            660                 665                 670
Ala Leu Pro Ala His Val Phe Arg Val Arg Arg His Ala Arg Arg Glu
        675                 680                 685
Gln Asp Phe Asp Tyr Phe Leu
    690                 695
```

<210> 414
<211> 2262

<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 414
atgtcgatgc agatacgtac agacaaggct caggagatcg atcgcaatgt tgtgctcact 60
atgcctgtca aggcatacga gactgtaacg gtcactaagg atggcctttc ttctttccct 120
gagattccga atgcagatga accggcgact cttttctata aagcccctgc aggatctgtg 180
tttgacggtt atactaaaga cttatatgct catatcggag tcattgatgg cgagtggaag 240
catgtggttt cggcatggga tgtagatgtg gctaagtgca aatggcataa gacagcccag 300
gataatgtat ggtcacttac actcgagcct tctatccgtg agtggttcgc ttccggcaca 360
actccagtga atgtgatcgg tgtagtggta agagatgcaa agggagcagg aagccttcag 420
acagaagacc tgactcttaa tgtgtatgac cctgtatatg gaacaggtga tatgcctaca 480
ccgcttccgg ctggagtaaa acatggaata aatatcattg acaacaagac cgtaactctt 540
gtgcttcttg agaaagataa gaatggtggt cactatgatg agtgctacct tatcggaact 600
gactttggca aatggagcag ggatgaaaaa tacaagatga acagggatga ggctgccggc 660
tgctggtgga tcactcttac cggtcttacc aggacaaagg aatatatgtt ccagtacgaa 720
cttgtcaagg gcaattcttc agttagggtg catgatccat attctgagat tgtttatgat 780
atgtacaacg atcagtggac atacgcaagc aataagatac ctaccggaaa gacagacggg 840
cttgtatctt cattccagat cagaaagtct acatattcat gggattaccc tgattttgag 900
attgaagatg aggatgacct tgtgatttat gaaatgcttg taagagactt caatacaagc 960
catagcatta gcgatgccat cgaccagctg gattatatct caggtcttgg agtgaatgct 1020
atagaactca tgcctatcca ggagtttgag gcaaacgaga gctggggata tgctcctcat 1080
tcatatttcg cattggataa atattacggt caacctacag aatataagaa attcgtcgat 1140
gcatgtcatg cgaggggtat tgcagtgatt gtagatgtag tatacaatca tgctacaggt 1200
gctcatccta tggccaagat gtattggagc ggtaatgcta catctgcatc caatccttgg 1260
ttcaatgtga cggcacctca tggtgacagt gtatatcatg actggaatca tagtaatgtc 1320
gatgttagaa atcatatcaa gagaagcctt gagtatctca tttctgaata taaggtggat 1380
ggattcagat ttgacctttc aaaaggtttc atccagagtg gaaatgagta tgactggaat 1440
caggacagaa caaactggat aaaggaatat tacaatacca taaaggcagc tgatccgaat 1500
gcagtggtta ttcttgagca ttgggtggat aatgagaatt atgacctctg caattatgga 1560
atgaaggtgt ggaacaaggc atgtgaacag tattatcagt ccggtatggg ttactcttcc 1620
gacagtgatt tctcaggaat cagagagcct tcatggttgc ctttcggttc atatatatca 1680
ttcatggaga gccatgatga agaaagagtc ggcttcaagc agaaagagta tggtaacggc 1740
agtgtcaaga cagatcttgc tttgagaatg cgtcgtgcag gccttaatgc agcattcttc 1800
cttactgttc caggtcctaa gatgatatgg cagttcggtg agatcggtta tgatatcagt 1860
atcgatcaca acggacgtac cggaaacaag cctgtcgtga ctgacgagta tctcgcaaac 1920
gagcatagga aggggcttta tgatacgtat gccggtcttt tggagttcag aaagaacaat 1980
ccgcgtttct ttgatgatga ttccaacttc cgttggggcg tttctacatc aaattggcca 2040
ggacgttata tcatgaatac gtcaaaggag ggtaagactt atgccatctt cggtaatttc 2100
ggctctggaa gtcagacaat aacgatggaa ctcccttctg aaggcccttg gtataactat 2160
tacaacagca ctgaggtgtg gaacggaaag aaccataaac cgactctcaa agaaggtgaa 2220
tttgtcttcc tcgtagacga taagagtctg tgcaagaatt ga 2262

<210> 415
<211> 753
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 415
Met Ser Met Gln Ile Arg Thr Asp Lys Ala Gln Glu Ile Asp Arg Asn
1 5 10 15
Val Val Leu Thr Met Pro Val Lys Ala Tyr Glu Thr Val Thr Val Thr
20 25 30
Lys Asp Gly Leu Ser Ser Phe Pro Glu Ile Pro Asn Ala Asp Glu Pro
35 40 45
Ala Thr Leu Phe Tyr Lys Ala Pro Ala Gly Ser Val Phe Asp Gly Tyr
50 55 60
Thr Lys Asp Leu Tyr Ala His Ile Gly Val Ile Asp Gly Glu Trp Lys
65 70 75 80
His Val Val Ser Ala Trp Asp Val Asp Val Ala Lys Cys Lys Trp His
85 90 95
Lys Thr Ala Gln Asp Asn Val Trp Ser Leu Thr Leu Glu Pro Ser Ile

```
                100                 105                 110
Arg Glu Trp Phe Ala Ser Gly Thr Thr Pro Val Asn Val Ile Gly Val
            115                 120                 125
Val Val Arg Asp Ala Lys Gly Ala Gly Ser Leu Gln Thr Glu Asp Leu
    130                 135                 140
Thr Leu Asn Val Tyr Asp Pro Val Tyr Gly Thr Gly Asp Met Pro Thr
145                 150                 155                 160
Pro Leu Pro Ala Gly Val Lys His Gly Ile Asn Ile Ile Asp Asn Lys
                165                 170                 175
Thr Val Thr Leu Val Leu Leu Glu Lys Asp Lys Asn Gly Gly His Tyr
            180                 185                 190
Asp Glu Cys Tyr Leu Ile Gly Thr Asp Phe Gly Lys Trp Ser Arg Asp
        195                 200                 205
Glu Lys Tyr Lys Met Asn Arg Asp Glu Ala Ala Gly Cys Trp Trp Ile
    210                 215                 220
Thr Leu Thr Gly Leu Thr Arg Thr Lys Glu Tyr Met Phe Gln Tyr Glu
225                 230                 235                 240
Leu Val Lys Gly Asn Ser Ser Val Arg Val His Asp Pro Tyr Ser Glu
                245                 250                 255
Ile Val Tyr Asp Met Tyr Asn Asp Gln Trp Thr Tyr Ala Ser Asn Lys
            260                 265                 270
Ile Pro Thr Gly Lys Thr Asp Gly Leu Val Ser Ser Phe Gln Ile Arg
        275                 280                 285
Lys Ser Thr Tyr Ser Trp Asp Tyr Pro Asp Phe Glu Ile Glu Asp Glu
    290                 295                 300
Asp Asp Leu Val Ile Tyr Glu Met Leu Val Arg Asp Phe Asn Thr Ser
305                 310                 315                 320
His Ser Ile Ser Asp Ala Ile Asp Gln Leu Asp Tyr Ile Ser Gly Leu
                325                 330                 335
Gly Val Asn Ala Ile Glu Leu Met Pro Ile Gln Glu Phe Glu Ala Asn
            340                 345                 350
Glu Ser Trp Gly Tyr Ala Pro His Ser Tyr Phe Ala Leu Asp Lys Tyr
        355                 360                 365
Tyr Gly Gln Pro Thr Glu Tyr Lys Lys Phe Val Asp Ala Cys His Ala
    370                 375                 380
Arg Gly Ile Ala Val Ile Val Asp Val Val Tyr Asn His Ala Thr Gly
385                 390                 395                 400
Ala His Pro Met Ala Lys Met Tyr Trp Ser Gly Asn Ala Thr Ser Ala
                405                 410                 415
Ser Asn Pro Trp Phe Asn Val Thr Ala Pro His Gly Asp Ser Val Tyr
            420                 425                 430
His Asp Trp Asn His Ser Asn Val Asp Val Arg Asn His Ile Lys Arg
        435                 440                 445
Ser Leu Glu Tyr Leu Ile Ser Glu Tyr Lys Val Asp Gly Phe Arg Phe
    450                 455                 460
Asp Leu Ser Lys Gly Phe Ile Gln Ser Gly Asn Glu Tyr Asp Trp Asn
465                 470                 475                 480
Gln Asp Arg Thr Asn Trp Ile Lys Glu Tyr Tyr Asn Thr Ile Lys Ala
                485                 490                 495
Ala Asp Pro Asn Ala Val Val Ile Leu Glu His Trp Val Asp Asn Glu
            500                 505                 510
Asn Tyr Asp Leu Cys Asn Tyr Gly Met Lys Val Trp Asn Lys Ala Cys
        515                 520                 525
Glu Gln Tyr Tyr Gln Ser Gly Met Gly Tyr Ser Ser Asp Ser Asp Phe
    530                 535                 540
Ser Gly Ile Arg Glu Pro Ser Trp Leu Pro Phe Gly Ser Tyr Ile Ser
545                 550                 555                 560
Phe Met Glu Ser His Asp Glu Glu Arg Val Gly Phe Lys Gln Lys Glu
                565                 570                 575
Tyr Gly Asn Gly Ser Val Lys Thr Asp Leu Ala Leu Arg Met Arg Arg
            580                 585                 590
Ala Gly Leu Asn Ala Ala Phe Phe Leu Thr Val Pro Gly Pro Lys Met
        595                 600                 605
Ile Trp Gln Phe Gly Glu Ile Gly Tyr Asp Ile Ser Ile Asp His Asn
    610                 615                 620
Gly Arg Thr Gly Asn Lys Pro Val Val Thr Asp Glu Tyr Leu Ala Asn
625                 630                 635                 640
Glu His Arg Lys Gly Leu Tyr Asp Thr Tyr Ala Gly Leu Leu Glu Phe
```

645 650 655
Arg Lys Asn Asn Pro Arg Phe Phe Asp Asp Asp Ser Asn Phe Arg Trp
660 665 670
Gly Val Ser Thr Ser Asn Trp Pro Gly Arg Tyr Ile Met Asn Thr Ser
675 680 685
Lys Glu Gly Lys Thr Tyr Ala Ile Phe Gly Asn Phe Gly Ser Gly Ser
690 695 700
Gln Thr Ile Thr Met Glu Leu Pro Ser Glu Gly Pro Trp Tyr Asn Tyr
705 710 715 720
Tyr Asn Ser Thr Glu Val Trp Asn Gly Lys Asn His Lys Pro Thr Leu
725 730 735
Lys Glu Gly Glu Phe Val Phe Leu Val Asp Asp Lys Ser Leu Cys Lys
740 745 750
Asn

<210> 416
<211> 1344
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 416
atgaaccggg ttaagagact ttcgattttg gtcgttcttt tccttaccgc cttcattccg 60
actgtttttg ccggcgagca gcctctcgcc atctttcacg cctttaacga ccccttcact 120
cttgttgaat cctatgtctg cgaactcgcc gggcagggat actcacatgt ccagatatct 180
ccagcgcaga agtcgaaccc tgctcgagcc tggtatgccc ggtatcaacc cgtagatttt 240
actgtcatcg aagggatggg cactgagagc gatctgagga agctcacgga taaggcccac 300
gcgtgtggaa taaaggtgat cgccgatgtg gtcttcaacc acatgtcgag catggacgag 360
tacaaggggc ttgacaagtt tccgggactt gctcctgctg atttccaccg gcagtgcggc 420
atcgattatt caaaacgaga ttcggtgcgg aactgttggc tcggaggcga cttgcccgat 480
ctggaccagt cccggccgag ggtacaggat gttcagagag cccacataag gaagctcctt 540
tccctcggca tagacggctt ccgcttcgat gcggctaaac acatcgaccc cattgttgtg 600
aaagactaca tcgatctcat cgacagggag agcaacggca ggacctggaa ctacctcgag 660
gtcatcgagg atgacggcac tcaggccacg gactacaact ggatagcggc agtgaccgat 720
ttcgtcctct acaaggagtc gttgaggaag gccttcagtc tcggcgggga cctgcgatcg 780
ctcaagatgc ctgtggctgt caatgattcg cggagtatcg tcttcgggag aaatcacgac 840
accgtgccgg agaataacca gaactgcatc gtcggctgct acgacagccg ggaggactcc 900
tatcttgcca cggcatacgt cctggcccgc gaatcgggag tcccgctggt cctcaactgg 960
gacaactacg acgcgcccta catcagcacc ggcgtgaagt tccgccagat catgacgcag 1020
cgaggacgat cggccatgaa cgtgaaggag aatgtgctgg gcgtcatcga cagtcctgtc 1080
gtcatgatga tggagcgcgg gagtgaaggc tttttcgtcc tcaacaagag cgccgaccgg 1140
ttcgatatcc cagttctgga tctgacactg accaatctcg agggatgtta tcgggagctg 1200
agaagaaaat tcaccgtcgc catcgagaga aagtacggta agaaatttgt cacccggtgg 1260
ggacgatggg accgggggg cctcgaaatc tacggccgcg acgctctcta cttcatacgg 1320
gaaccctggg agcagtgcag gtaa 1344

<210> 417
<211> 447
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(24)

<400> 417
Met Asn Arg Val Lys Arg Leu Ser Ile Leu Val Val Leu Phe Leu Thr
1 5 10 15
Ala Phe Ile Pro Thr Val Phe Ala Gly Glu Gln Pro Leu Ala Ile Phe
20 25 30
His Ala Phe Asn Asp Pro Phe Thr Leu Val Glu Ser Tyr Val Cys Glu
35 40 45
Leu Ala Gly Gln Gly Tyr Ser His Val Gln Ile Ser Pro Ala Gln Lys

```
    50                  55                  60
Ser Asn Pro Ala Arg Ala Trp Tyr Ala Arg Tyr Gln Pro Val Asp Phe
65                  70                  75                  80
Thr Val Ile Glu Gly Met Gly Thr Glu Ser Asp Leu Arg Lys Leu Thr
                85                  90                  95
Asp Lys Ala His Ala Cys Gly Ile Lys Val Ile Ala Asp Val Val Phe
            100                 105                 110
Asn His Met Ser Ser Met Asp Glu Tyr Lys Gly Leu Asp Lys Phe Pro
        115                 120                 125
Gly Leu Ala Pro Ala Asp Phe His Arg Gln Cys Gly Ile Asp Tyr Ser
    130                 135                 140
Lys Arg Asp Ser Val Arg Asn Cys Trp Leu Gly Gly Asp Leu Pro Asp
145                 150                 155                 160
Leu Asp Gln Ser Arg Pro Arg Val Gln Asp Val Gln Arg Ala His Ile
                165                 170                 175
Arg Lys Leu Leu Ser Leu Gly Ile Asp Gly Phe Arg Phe Asp Ala Ala
            180                 185                 190
Lys His Ile Asp Pro Ile Val Val Lys Asp Tyr Ile Asp Leu Ile Asp
        195                 200                 205
Arg Glu Ser Asn Gly Arg Thr Trp Asn Tyr Leu Glu Val Ile Glu Asp
    210                 215                 220
Asp Gly Thr Gln Ala Thr Asp Tyr Asn Trp Ile Ala Ala Val Thr Asp
225                 230                 235                 240
Phe Val Leu Tyr Lys Glu Ser Leu Arg Lys Ala Phe Ser Leu Gly Gly
                245                 250                 255
Asp Leu Arg Ser Leu Lys Met Pro Val Ala Val Asn Asp Ser Arg Ser
            260                 265                 270
Ile Val Phe Gly Arg Asn His Asp Thr Val Pro Glu Asn Asn Gln Asn
        275                 280                 285
Cys Ile Val Gly Cys Tyr Asp Ser Arg Glu Asp Ser Tyr Leu Ala Thr
    290                 295                 300
Ala Tyr Val Leu Ala Arg Glu Ser Gly Val Pro Leu Val Leu Asn Trp
305                 310                 315                 320
Asp Asn Tyr Asp Ala Pro Tyr Ile Ser Thr Gly Val Lys Phe Arg Gln
                325                 330                 335
Ile Met Thr Gln Arg Gly Arg Ser Ala Met Asn Val Lys Glu Asn Val
            340                 345                 350
Leu Gly Val Ile Asp Ser Pro Val Val Met Met Met Glu Arg Gly Ser
        355                 360                 365
Glu Gly Phe Phe Val Leu Asn Lys Ser Ala Asp Arg Phe Asp Ile Pro
    370                 375                 380
Val Leu Asp Leu Thr Leu Thr Asn Leu Glu Gly Cys Tyr Arg Glu Leu
385                 390                 395                 400
Arg Arg Lys Phe Thr Val Ala Ile Glu Arg Lys Tyr Gly Lys Lys Phe
                405                 410                 415
Val Thr Arg Trp Gly Arg Trp Asp Arg Gly Gly Leu Glu Ile Tyr Gly
            420                 425                 430
Arg Asp Ala Leu Tyr Phe Ile Arg Glu Pro Trp Glu Gln Cys Arg
        435                 440                 445
```

<210> 418
<211> 1455
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 418

```
atgaaccgtc caggcacggg cgcctcgggg cgcccacaat ctcgctccgc cacgtcgtgg      60
caatcccgca acggcggctg gctgctcgcc tcgctgctgg ccgtgtgttt cgcaacggcg     120
cccgtgcgcg ccgatgtcat cctgcatgca ttcaactggc cgtatgcgac ggtcgaagcc     180
cgcgccaacg agctgcgaga cctgggctat cgcgccgtac tggtcgcacc gccggtgaag     240
tccgaaggca atgcgtggtg ggcgcgctac cagccgcagg actaccgggt catcgagcat     300
ccgctcggca atcgcgagtc gttcgtgcgc atgtccagtg ctctgcgcgc ccgcggcatt     360
cgcgtctacg ccgacatcgt gctgaatcac atggccaatg aagccccgca gcgaccggat     420
ctgaactacc ccggtcaacg ggtcctggat cagtatgcag gcaataccgc gtacttcgcg     480
cagcaacgac tctacggcga tctccgctac aacttcatgt cggcctggga ctttggcccc     540
```

```
gcccactgca tcggcaacta ccacgacgtc tggcaggtgc agaactggcg gctgtgcagc     600
ggcgctggcg atgccggcct gcccgacctg ctggccagcg attacatcgt cggacagcag     660
cgcacctatc tccaggcact gaagaacctc ggcgtctccg ggttgcgcat cgacgcggcc     720
aagcacatgc cgctcagcca catcaaccgc gtactgactg ccgacctcaa ggccggcatg     780
catgtgttcg gcgaagtcat cacccacggt ggcgtcggtg atcccgaata cgatctgttt     840
ctccggccct atctggatgg caccgatcac ggtgcctacg acttcccgct cttcgaggcg     900
atccgccgcg cattcggctt cggtggcagc atgagcaccc tggtcgatcc tggggccgtc     960
ggactggccc tgcccaatgc ccggtcgatc accttcaccg tcacccacga catcccgaac    1020
aatggtgtgt tccggcatct gctgctggat gccggcgacg aaaccctggc gtacgcctac    1080
atcctgggtc gtgacggcgg cagtccactg ctgtactccg atcacaacga gagcggcgac    1140
aaccgctggg ttcatgccta tcggcgcaac gacctcgcgg ccatgatccg ctttcacaat    1200
gccaatcatg gcaatgacat gcaggtgctc gcacacggga actgccatct gctctttcgc    1260
cgcggcaatc gcggcatcgt ggcgatcaac aagtgcgggc atacggtcaa cgccacggtg    1320
aacatgaaca acagtgtgct gtggtggcat accccgtatc gcgatgtcct tgatgccggc    1380
agcgtggtgc agattggcag cgccagccac acgttcagcc ttccgccgcg gcgcgctcgg    1440
atgtggctgc gctga                                                     1455

<210> 419
<211> 484
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(44)

<400> 419
Met Asn Arg Pro Gly Thr Gly Ala Ser Gly Arg Pro Gln Ser Arg Ser
 1               5                  10                  15
Ala Thr Ser Trp Gln Ser Arg Asn Gly Gly Trp Leu Leu Ala Ser Leu
            20                  25                  30
Leu Ala Val Cys Phe Ala Thr Ala Pro Val Arg Ala Asp Val Ile Leu
        35                  40                  45
His Ala Phe Asn Trp Pro Tyr Ala Thr Val Glu Ala Arg Ala Asn Glu
    50                  55                  60
Leu Arg Asp Leu Gly Tyr Arg Ala Val Leu Val Ala Pro Pro Val Lys
65                  70                  75                  80
Ser Glu Gly Asn Ala Trp Trp Ala Arg Tyr Gln Pro Gln Asp Tyr Arg
                85                  90                  95
Val Ile Glu His Pro Leu Gly Asn Arg Glu Ser Phe Val Arg Met Ser
            100                 105                 110
Ser Ala Leu Arg Ala Arg Gly Ile Arg Val Tyr Ala Asp Ile Val Leu
        115                 120                 125
Asn His Met Ala Asn Glu Ala Pro Gln Arg Pro Asp Leu Asn Tyr Pro
    130                 135                 140
Gly Gln Arg Val Leu Asp Gln Tyr Ala Gly Asn Thr Ala Tyr Phe Ala
145                 150                 155                 160
Gln Gln Arg Leu Tyr Gly Asp Leu Arg Tyr Asn Phe Met Ser Ala Trp
                165                 170                 175
Asp Phe Gly Pro Ala His Cys Ile Gly Asn Tyr His Asp Val Trp Gln
            180                 185                 190
Val Gln Asn Trp Arg Leu Cys Ser Gly Ala Gly Asp Ala Gly Leu Pro
        195                 200                 205
Asp Leu Leu Ala Ser Asp Tyr Ile Val Gly Gln Gln Arg Thr Tyr Leu
    210                 215                 220
Gln Ala Leu Lys Asn Leu Gly Val Ser Gly Leu Arg Ile Asp Ala Ala
225                 230                 235                 240
Lys His Met Pro Leu Ser His Ile Asn Arg Val Leu Thr Ala Asp Leu
                245                 250                 255
Lys Ala Gly Met His Val Phe Gly Glu Val Ile Thr His Gly Gly Val
            260                 265                 270
Gly Asp Pro Glu Tyr Asp Leu Phe Leu Arg Pro Tyr Leu Asp Gly Thr
        275                 280                 285
Asp His Gly Ala Tyr Asp Phe Pro Leu Phe Glu Ala Ile Arg Arg Ala
    290                 295                 300
Phe Gly Phe Gly Gly Ser Met Ser Thr Leu Val Asp Pro Gly Ala Val
```

305 310 315 320
Gly Leu Ala Leu Pro Asn Ala Arg Ser Ile Thr Phe Thr Val Thr His
325 330 335
Asp Ile Pro Asn Asn Gly Val Phe Arg His Leu Leu Leu Asp Ala Gly
340 345 350
Asp Glu Thr Leu Ala Tyr Ala Tyr Ile Leu Gly Arg Asp Gly Gly Ser
355 360 365
Pro Leu Leu Tyr Ser Asp His Asn Glu Ser Gly Asp Asn Arg Trp Val
370 375 380
His Ala Tyr Arg Arg Asn Asp Leu Ala Ala Met Ile Arg Phe His Asn
385 390 395 400
Ala Asn His Gly Asn Asp Met Gln Val Leu Ala His Gly Asn Cys His
405 410 415
Leu Leu Phe Arg Arg Gly Asn Arg Gly Ile Val Ala Ile Asn Lys Cys
420 425 430
Gly His Thr Val Asn Ala Thr Val Asn Met Asn Asn Ser Val Leu Trp
435 440 445
Trp His Thr Pro Tyr Arg Asp Val Leu Asp Ala Gly Ser Val Val Gln
450 455 460
Ile Gly Ser Ala Ser His Thr Phe Ser Leu Pro Pro Arg Arg Ala Arg
465 470 475 480
Met Trp Leu Arg

<210> 420
<211> 1308
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 420
atgcgttcat cagtgttgtg cggaggggcc tggctgtatt gggcccgtcg tttcgcttcg 60
aacaagccag tcatcacccg tcgtcttggc gccgcacttg ccggtctgct gttttgtgcc 120
ggtgctcacg cggcgaccga acagagcaag gccagtagcg caatcctgtt ccagggcttt 180
cactggcact ccgccaatgg cagctggtac ggcaacctgc agggcaaggc cggtgatctg 240
aaggatctcg gcatcaccca tgtctggttc ccgccgccgt cggacgccgc gtcgctcgaa 300
ggctacctgc cgcgccagct taatgtgctg aactccagtt acggcaacga aaccgcgctg 360
aagaacgcga ttagcgcgct gaacggccag ggcatcaaga gcgtggttga cgtcgtgatc 420
aaccatcgag tgggtacagc caactgggcc gatttcacca acccgacctg ggattgccgc 480
gccgtcacca gcaatgatga atggagcggc cgctgtggca ccgccgacac cggcgatggt 540
tacgatgctg cgcgcgatct cgaccaccgc cagaccttcc tgcagaacga tctcaagact 600
tggctgtcgt cgcgtctcaa aggtgtgggc ttcagcggca ttcgctacga ctactcgaaa 660
ggctacggac cgggttacgc cgggctttac cacgatgcga tggcgccgaa cttctgcgtc 720
ggcgagatct ggaccaatct tgattacaac aatgtcgatg cccatcgcca gctgctgatg 780
aactacgtcg atggcaacgg cggcaagtgt ggcgcgttcg acttcaccac caagggcctg 840
ctgaatcagg cgctgtcggc caacgaatac tggcgcctgc gcgcttccga tggcaaaccg 900
gccggcggta tcggctggtg ggcacagaag atggtgacct ttgtcgacaa ccacgacacc 960
gggccgtcgc agagttgcgg cagcggccaa aatcactggc cggtgccctg cggcaaggtg 1020
atgcagggct atgcctacgt gctgagccat ccgggtgttc cgactgtgta ctacccgcat 1080
gtctacgact ggaacctgcg tgccgcgatc aagacgctga tccagctgcg caaggaaaaa 1140
ggcatcacct cgacctcgtc ggttgcaatc cagcgtgccg atcagggcct ttacgcggcg 1200
atcatcaaca acaacctggc gatgaagatc ggcccgaact catggagtcc tggcagcggc 1260
tggacactgc gcacttccgg cgatcaatac gcgatctgga cccgctga 1308

<210> 421
<211> 435
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(44)

<400> 421

```
Met Arg Ser Ser Val Leu Cys Gly Gly Ala Trp Leu Tyr Trp Ala Arg
 1               5                  10                  15
Arg Phe Ala Ser Asn Lys Pro Val Ile Thr Arg Arg Leu Gly Ala Ala
            20                  25                  30
Leu Ala Gly Leu Leu Phe Cys Ala Gly Ala His Ala Ala Thr Glu Gln
        35                  40                  45
Ser Lys Ala Ser Ser Ala Ile Leu Phe Gln Gly Phe His Trp His Ser
    50                  55                  60
Ala Asn Gly Ser Trp Tyr Gly Asn Leu Gln Gly Lys Ala Gly Asp Leu
65                  70                  75                  80
Lys Asp Leu Gly Ile Thr His Val Trp Phe Pro Pro Pro Ser Asp Ala
                85                  90                  95
Ala Ser Leu Glu Gly Tyr Leu Pro Arg Gln Leu Asn Val Leu Asn Ser
            100                 105                 110
Ser Tyr Gly Asn Glu Thr Ala Leu Lys Asn Ala Ile Ser Ala Leu Asn
        115                 120                 125
Gly Gln Gly Ile Lys Ser Val Val Asp Val Val Ile Asn His Arg Val
    130                 135                 140
Gly Thr Ala Asn Trp Ala Asp Phe Thr Asn Pro Thr Trp Asp Cys Arg
145                 150                 155                 160
Ala Val Thr Ser Asn Asp Glu Trp Ser Gly Arg Cys Gly Thr Ala Asp
                165                 170                 175
Thr Gly Asp Gly Tyr Asp Ala Ala Arg Asp Leu Asp His Arg Gln Thr
            180                 185                 190
Phe Leu Gln Asn Asp Leu Lys Thr Trp Leu Ser Ser Arg Leu Lys Gly
        195                 200                 205
Val Gly Phe Ser Gly Ile Arg Tyr Asp Tyr Ser Lys Gly Tyr Gly Pro
    210                 215                 220
Gly Tyr Ala Gly Leu Tyr His Asp Ala Met Ala Pro Asn Phe Cys Val
225                 230                 235                 240
Gly Glu Ile Trp Thr Asn Leu Asp Tyr Asn Asn Val Asp Ala His Arg
                245                 250                 255
Gln Leu Leu Met Asn Tyr Val Asp Gly Asn Gly Gly Lys Cys Gly Ala
            260                 265                 270
Phe Asp Phe Thr Thr Lys Gly Leu Leu Asn Gln Ala Leu Ser Ala Asn
        275                 280                 285
Glu Tyr Trp Arg Leu Arg Ala Ser Asp Gly Lys Pro Ala Gly Gly Ile
    290                 295                 300
Gly Trp Trp Ala Gln Lys Met Val Thr Phe Val Asp Asn His Asp Thr
305                 310                 315                 320
Gly Pro Ser Gln Ser Cys Gly Ser Gly Gln Asn His Trp Pro Val Pro
                325                 330                 335
Cys Gly Lys Val Met Gln Gly Tyr Ala Tyr Val Leu Ser His Pro Gly
            340                 345                 350
Val Pro Thr Val Tyr Tyr Pro His Val Tyr Asp Trp Asn Leu Arg Ala
        355                 360                 365
Ala Ile Lys Thr Leu Ile Gln Leu Arg Lys Glu Lys Gly Ile Thr Ser
    370                 375                 380
Thr Ser Ser Val Ala Ile Gln Arg Ala Asp Gln Gly Leu Tyr Ala Ala
385                 390                 395                 400
Ile Ile Asn Asn Asn Leu Ala Met Lys Ile Gly Pro Asn Ser Trp Ser
                405                 410                 415
Pro Gly Ser Gly Trp Thr Leu Arg Thr Ser Gly Asp Gln Tyr Ala Ile
            420                 425                 430
Trp Thr Arg
        435

<210> 422
<211> 2751
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 422
atgaaaacct cctactcctt tgtatataca atacttatcg tcttcgtact tattccgctt      60
gcacctgttc gggtaagtgc acaagtagat tttcgcgaag aaacgattta tttcctgctc     120
```

```
acgacacgct tttttgacgg cgattcgtcg aacaatgtac ccaacgaatg gtcgagctac     180
caccccgacc cgaatatcaa tccatcgatt acagaccctc aggatgttac gtggcggggt     240
gactttaaag gacttattca gaaactcgac tacattcagg accttggatt taccgcaata     300
tggataaccc ctatcgtgca taatcgcagc ccgctggact accacggcta ccatgcctgg     360
gactttacac gcgtggatcc gcgcctggaa tcacccggag ccacctttca ggacctcatc     420
aacgctgtac atgcccgggg gatgaaaatc gtgctggata ttgtaaccaa tcactccggc     480
cggtttggaa ttaaggattt tgccgaaatc aagtataata ccgatcccaa ccgcccatgg     540
ggacaagacg cacagggcaa ccccctgcag gacaacccca actgggaata cgacggactt     600
accccaaacc ccgatgatgg caaaatctgg agtcgtgcca acctggccag gttacctgcc     660
ccttataacg aaaacctggc ggcctacaac ttcccaaaca ccataagcta cgtgaatacc     720
accagcccag actggtttca tcattccgga aacgggtttg cccagggctt cgacgatatc     780
gaaaatctgc acaaccgcgc cctgcacgaa gacacgcccg acctcaatac cagttccccc     840
caggtgcgtc aatacttgct cgacgcctac cgtaccttta ttgaaatggg tgtggacggt     900
ttccgctggg acaccattaa gcatatgagc cgcagcgatg tcctgtggtt tctggacgaa     960
ttcaaaaaga ttaaccccga cctgttcatt ttcggcgaag tagcccaaaa acggcacgag    1020
ctgcacaatg tagaagaaac caatccacat tggtacacat ggagaggaga ccccggcaac    1080
tccgaaccct ccgggatgtc cgttatcgac ttttttgcca tggccacctt ccacttgttc    1140
gaaaaaggcg agaatttttc gaatgtgatg gcggccgccc gctacgacca tctgtatgcc    1200
aaccccggcg aactggtcac gtttttggat aaccacgact tcgggcccaa caacgactgg    1260
aaccgccgct ttggcggaga tgcccgcaac ctcactgccg ccctcaactt tatgtggacc    1320
tggcggggaa ttccttccct ttattatggc acagaagtcc aatttatgcg tggcgccttt    1380
accgacatac agtatggagg acatcagcgc tccatcgacg aaaccggcag agcttatttt    1440
ggcgaacagt tcgatgatgc acccaatcat atcgtctacc agcatatccg taaactgaat    1500
gccatgcgac gcgctgttcc tgccctgcaa aaaggctcct ggcgctgggg cggcaatggt    1560
ggtgccaacg gcgtcggcat cgtaagggaa tacggcgata gctatgctgt agtcggtcta    1620
gccaaagacg gagcagtaac cttcaacttc accaatattc gaaacggaac ctaccgggat    1680
gccgttaccg gccacgaaat atcggtctcc aataactcca tccgctttac cgtagcccct    1740
tcctctgcag gaatttatgt actcaacgga cccggccgca tcggcgatct cggtgccggg    1800
ttcttccagt ccgactccga gggaagtcag ggcgatggcg gagataatgg aggcggtggc    1860
ggaaatcagg aaccggccag cgtctccttc aacccaaatc cacccgttat tggtcagcag    1920
gtaaccttaa cctacgatgg tttcctgaac tcccagcaac aggttaatct gcactggggc    1980
tacgacggct ggacaaacgt atccaccacg cctatgcaga aagacggcga cgtctggaaa    2040
ctgaccttca cgataccatc gaccgcgacc tcccgactaa atttcgtatt caataacgga    2100
aacgatgtct gggataacaa caacagcaac aactattcgg taacgctggc cagtcccgac    2160
ggcagcaccc cgcctcccgt atttaccgcg aatccggatc ctgccgtggc gggagaaccc    2220
ctcaccctga cctatcaggg ctcgctggcc ggtgcatcgg ctgttaacct gcactggggc    2280
tacgacggct ggacgaacgt aaccaatacc cccatgagca aaagtggtga tgtgtggacc    2340
ctcaccatga ccgttcccgc cactgctacc cgcgagctga atttcgtatt caacgacggg    2400
aacaacacct gggacaacaa cggaggccag gactacctgc tcaacgtaac caccaccacc    2460
tcagccgact tcgaagaccc gttccccaac gccttccggc tgctgggcaa ctaccccaat    2520
ccatttaacc ccactaccag catcagcttc gaaatacccg aaggaggcac cgtgcggatt    2580
cagatcttcg acgtatccgg caggcttatc gaagagctct ctgcggccta tccgggcgcc    2640
ggtgtgcatg ccctgcgctt cgatgcacgc caccttttcta ccggtatgta cctgtaccga    2700
gccactttcg gcggacacgt actaactggc aaaatgacac tcgtcaagta a              2751

<210> 423
<211> 916
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(27)

<400> 423
Met Lys Thr Ser Tyr Ser Phe Val Tyr Thr Ile Leu Ile Val Phe Val
 1               5                  10                  15
Leu Ile Pro Leu Ala Pro Val Arg Val Ser Ala Gln Val Asp Phe Arg
            20                  25                  30
Glu Glu Thr Ile Tyr Phe Leu Leu Thr Thr Arg Phe Phe Asp Gly Asp
        35                  40                  45
Ser Ser Asn Asn Val Pro Asn Glu Trp Ser Ser Tyr His Pro Asp Pro
    50                  55                  60
Asn Ile Asn Pro Ser Ile Thr Asp Pro Gln Asp Val Thr Trp Arg Gly
65                  70                  75                  80
Asp Phe Lys Gly Leu Ile Gln Lys Leu Asp Tyr Ile Gln Asp Leu Gly
```

```
                85                  90                  95
Phe Thr Ala Ile Trp Ile Thr Pro Ile Val His Asn Arg Ser Pro Leu
            100                 105                 110
Asp Tyr His Gly Tyr His Ala Trp Asp Phe Thr Arg Val Asp Pro Arg
        115                 120                 125
Leu Glu Ser Pro Gly Ala Thr Phe Gln Asp Leu Ile Asn Ala Val His
    130                 135                 140
Ala Arg Gly Met Lys Ile Val Leu Asp Ile Val Thr Asn His Ser Gly
145                 150                 155                 160
Arg Phe Gly Ile Lys Asp Phe Ala Glu Ile Lys Tyr Asn Thr Asp Pro
                165                 170                 175
Asn Arg Pro Trp Gly Gln Asp Ala Gln Gly Asn Pro Leu Gln Asp Asn
            180                 185                 190
Pro Asn Trp Glu Tyr Asp Gly Leu Thr Pro Asn Pro Asp Asp Gly Lys
        195                 200                 205
Ile Trp Ser Arg Ala Asn Leu Ala Arg Leu Pro Ala Pro Tyr Asn Glu
    210                 215                 220
Asn Leu Ala Ala Tyr Asn Phe Pro Asn Thr Ile Ser Tyr Val Asn Thr
225                 230                 235                 240
Thr Ser Pro Asp Trp Phe His His Ser Gly Asn Gly Phe Ala Gln Gly
                245                 250                 255
Phe Asp Asp Ile Glu Asn Leu His Asn Arg Ala Leu His Glu Asp Thr
            260                 265                 270
Pro Asp Leu Asn Thr Ser Ser Pro Gln Val Arg Gln Tyr Leu Leu Asp
        275                 280                 285
Ala Tyr Arg Thr Phe Ile Glu Met Gly Val Asp Gly Phe Arg Trp Asp
    290                 295                 300
Thr Ile Lys His Met Ser Arg Ser Asp Val Leu Trp Phe Leu Asp Glu
305                 310                 315                 320
Phe Lys Lys Ile Asn Pro Asp Leu Phe Ile Phe Gly Glu Val Ala Gln
                325                 330                 335
Lys Arg His Glu Leu His Asn Val Glu Glu Thr Asn Pro His Trp Tyr
            340                 345                 350
Thr Trp Arg Gly Asp Pro Gly Asn Ser Glu Pro Ser Gly Met Ser Val
        355                 360                 365
Ile Asp Phe Phe Ala Met Ala Thr Phe His Leu Phe Glu Lys Gly Glu
    370                 375                 380
Asn Phe Ser Asn Val Met Ala Ala Ala Arg Tyr Asp His Leu Tyr Ala
385                 390                 395                 400
Asn Pro Gly Glu Leu Val Thr Phe Leu Asp Asn His Asp Phe Gly Pro
                405                 410                 415
Asn Asn Asp Trp Asn Arg Arg Phe Gly Gly Asp Ala Arg Asn Leu Thr
            420                 425                 430
Ala Ala Leu Asn Phe Met Trp Thr Trp Arg Gly Ile Pro Ser Leu Tyr
        435                 440                 445
Tyr Gly Thr Glu Val Gln Phe Met Arg Gly Ala Phe Thr Asp Ile Gln
    450                 455                 460
Tyr Gly Gly His Gln Arg Ser Ile Asp Glu Thr Gly Arg Ala Tyr Phe
465                 470                 475                 480
Gly Glu Gln Phe Asp Asp Ala Pro Asn His Ile Val Tyr Gln His Ile
                485                 490                 495
Arg Lys Leu Asn Ala Met Arg Arg Ala Val Pro Ala Leu Gln Lys Gly
            500                 505                 510
Ser Trp Arg Trp Gly Gly Asn Gly Gly Ala Asn Gly Val Gly Ile Val
        515                 520                 525
Arg Glu Tyr Gly Asp Ser Tyr Ala Val Val Gly Leu Ala Lys Asp Gly
    530                 535                 540
Ala Val Thr Phe Asn Phe Thr Asn Ile Arg Asn Gly Thr Tyr Arg Asp
545                 550                 555                 560
Ala Val Thr Gly His Glu Ile Ser Val Ser Asn Asn Ser Ile Arg Phe
                565                 570                 575
Thr Val Ala Pro Ser Ser Ala Gly Ile Tyr Val Leu Asn Gly Pro Gly
            580                 585                 590
Arg Ile Gly Asp Leu Gly Ala Gly Phe Phe Gln Ser Asp Ser Glu Gly
        595                 600                 605
Ser Gln Gly Asp Gly Gly Asp Asn Gly Gly Gly Gly Gly Asn Gln Glu
    610                 615                 620
Pro Ala Ser Val Ser Phe Asn Pro Asn Pro Pro Val Ile Gly Gln Gln
```

625 630 635 640
Val Thr Leu Thr Tyr Asp Gly Phe Leu Asn Ser Gln Gln Gln Val Asn
645 650 655
Leu His Trp Gly Tyr Asp Gly Trp Thr Asn Val Ser Thr Thr Pro Met
660 665 670
Gln Lys Asp Gly Asp Val Trp Lys Leu Thr Phe Thr Ile Pro Ser Thr
675 680 685
Ala Thr Ser Arg Leu Asn Phe Val Phe Asn Asn Gly Asn Asp Val Trp
690 695 700
Asp Asn Asn Asn Ser Asn Asn Tyr Ser Val Thr Leu Ala Ser Pro Asp
705 710 715 720
Gly Ser Thr Pro Pro Pro Val Phe Thr Ala Asn Pro Asp Pro Ala Val
725 730 735
Ala Gly Glu Pro Leu Thr Leu Thr Tyr Gln Gly Ser Leu Ala Gly Ala
740 745 750
Ser Ala Val Asn Leu His Trp Gly Tyr Asp Gly Trp Thr Asn Val Thr
755 760 765
Asn Thr Pro Met Ser Lys Ser Gly Asp Val Trp Thr Leu Thr Met Thr
770 775 780
Val Pro Ala Thr Ala Thr Arg Glu Leu Asn Phe Val Phe Asn Asp Gly
785 790 795 800
Asn Asn Thr Trp Asp Asn Asn Gly Gly Gln Asp Tyr Leu Leu Asn Val
805 810 815
Thr Thr Thr Thr Ser Ala Asp Phe Glu Asp Pro Phe Pro Asn Ala Phe
820 825 830
Arg Leu Leu Gly Asn Tyr Pro Asn Pro Phe Asn Pro Thr Thr Ser Ile
835 840 845
Ser Phe Glu Ile Pro Glu Gly Gly Thr Val Arg Ile Gln Ile Phe Asp
850 855 860
Val Ser Gly Arg Leu Ile Glu Glu Leu Ser Ala Ala Tyr Pro Gly Ala
865 870 875 880
Gly Val His Ala Leu Arg Phe Asp Ala Arg His Leu Ser Thr Gly Met
885 890 895
Tyr Leu Tyr Arg Ala Thr Phe Gly Gly His Val Leu Thr Gly Lys Met
900 905 910
Thr Leu Val Lys
915

<210> 424
<211> 2169
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 424
gtggattcta tgtttaagca ggtcaaacgt ctattcctta gcctggcgct taccctgggc 60
ttactcaccg gaagtgcgct tccgctagtc cccgtctccc aggcctttgc cgatcccgat 120
actgcggtat ccaataaggc cagcttcagc accgacgtca tttatcaagt ttttaccgac 180
cgctttctcg acggggatgc ctccaataac ccgaccggag cggcctatga ctccaccggc 240
accaatctga agctgtatct cggcggcgac tggaaagggc tgaccgataa gctgaatgat 300
aattatttta cggatctggg cattacggca ctctggatct cccagcctgt cgagaatatt 360
tattcggttg tcaactattc cggagtgaac agcacttcct accatggcta ctgggcccgc 420
gacttcaaga agaccaaccc ctatttcggc tcgatggccg atttccagac tctcgttaat 480
accgctcatg caaaaggcat caagatcgtc atcgactttg ctccgaatca tacctctcct 540
gcaatggaga ccgataccag cttcgccgag aacggcaaat tgtacgacaa cggcaccctg 600
ctgggcggct acacgaatga cacgaacgga tttttccatc acaacggcgg ctccgatttc 660
tccacactgg agaacggcat ttacaagaac ctttacgatc tggccgatct gaatcacaat 720
aacagcacga ttgacaagta cttcaaagat gcgatcaagg tgtggatcga tacgggcgtg 780
gacggcatcc gcgttgatgc agtcaagcac atgccggacg gctggcagaa aaactgggta 840
tccagcattt acaactacga gcccgtgttt attttcgggg aatggtacct cggttcgtcg 900
gctgccgatg cagataatac gaagttcgcc aacacctcgg ggatgagcct gcttgatttc 960
cggttcaacc aggaggtgcg gaacgtattc cgcaatgaca cctccaccat gtatgcactc 1020
gattccatgc tcaccaatac agctgctgac tacacccagg tcaacgatca agtaaccttt 1080
attgataatc atgatatgga ccgcttcaag accggcacgc tgaacaaccg ccgtctggag 1140
caggctctgg ccttcactct gacctcccgc ggcgttcctg ccatttatta cggaacagaa 1200
cagtatatga ccggcaacgg cgacccggac aaccgggcga tgatgacttc cttctccacc 1260

tctactaccg cgttcaatgt catcagcaag cttgcgccgc tgcgcaaatc caatccggcc 1320
atcgcttacg gcactaccca gcagcgctgg atcaacaatg acgtgtatat ctatgaacgc 1380
aaattcggca acaacgtcgc tgtcgtagcg gtcaaccgca atctcagcac agcaacatcc 1440
atcagcggac tggttacttc cttgccgtca ggaacctaca gcgatgtgct gggcggcact 1500
ttaagcggca attccattac cgcaagctcc ggtaacgtcg ccacctttac gctgggcgca 1560
ggcgcagcgg cggtatggca atacacggag acgagctcgg ccactcccgt gctcggccat 1620
gtcggcccga tgatggggca acccggcaat gaagtaacca tcgacggacg cggcttcggc 1680
tcgtctgcag gaacggtcta tttcggcaca accgccgtaa ccggctcgaa cattatttcg 1740
tgggaagata cacaaatcaa ggtgaaggta ccttccgtag cggccggcac ctacgcggtt 1800
aaagtaacga acagcgccgg aacaagcaat tcttacagcg gctttaccat tctcaccggc 1860
actcaggttt ccgtacgctt tatcgtgaac aacgccacta catcgcttgg ccaaaatgta 1920
tatctcaccg gcaatgtagc cgagctcggt aactggacta caggcagcag tgcaatcggg 1980
ccgttgttca accagatcat taaagtctac cctacctggt attatgatgt cagcgttccg 2040
gccggcactg cgctcgaatt caaatttttc aagaaaagcg gttcgaccgt tacctgggag 2100
agcggctcta atcatacgtt caccgctcct gcgagcggaa ccgcaactgt aaccgtggat 2160
tggcagtaa 2169

<210> 425
<211> 722
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(37)

<400> 425
Met Asp Ser Met Phe Lys Gln Val Lys Arg Leu Phe Leu Ser Leu Ala
 1   5   10   15
Leu Thr Leu Gly Leu Leu Thr Gly Ser Ala Leu Pro Leu Val Pro Val
            20   25   30
Ser Gln Ala Phe Ala Asp Pro Asp Thr Ala Val Ser Asn Lys Ala Ser
        35   40   45
Phe Ser Thr Asp Val Ile Tyr Gln Val Phe Thr Asp Arg Phe Leu Asp
    50   55   60
Gly Asp Ala Ser Asn Asn Pro Thr Gly Ala Ala Tyr Asp Ser Thr Gly
65   70   75   80
Thr Asn Leu Lys Leu Tyr Leu Gly Gly Asp Trp Lys Gly Leu Thr Asp
                85   90   95
Lys Leu Asn Asp Asn Tyr Phe Thr Asp Leu Gly Ile Thr Ala Leu Trp
            100   105   110
Ile Ser Gln Pro Val Glu Asn Ile Tyr Ser Val Val Asn Tyr Ser Gly
        115   120   125
Val Asn Ser Thr Ser Tyr His Gly Tyr Trp Ala Arg Asp Phe Lys Lys
    130   135   140
Thr Asn Pro Tyr Phe Gly Ser Met Ala Asp Phe Gln Thr Leu Val Asn
145   150   155   160
Thr Ala His Ala Lys Gly Ile Lys Ile Val Ile Asp Phe Ala Pro Asn
                165   170   175
His Thr Ser Pro Ala Met Glu Thr Asp Thr Ser Phe Ala Glu Asn Gly
            180   185   190
Lys Leu Tyr Asp Asn Gly Thr Leu Leu Gly Gly Tyr Thr Asn Asp Thr
        195   200   205
Asn Gly Phe Phe His His Asn Gly Gly Ser Asp Phe Ser Thr Leu Glu
    210   215   220
Asn Gly Ile Tyr Lys Asn Leu Tyr Asp Leu Ala Asp Leu Asn His Asn
225   230   235   240
Asn Ser Thr Ile Asp Lys Tyr Phe Lys Asp Ala Ile Lys Val Trp Ile
                245   250   255
Asp Thr Gly Val Asp Gly Ile Arg Val Asp Ala Val Lys His Met Pro
            260   265   270
Asp Gly Trp Gln Lys Asn Trp Val Ser Ser Ile Tyr Asn Tyr Glu Pro
        275   280   285
Val Phe Ile Phe Gly Glu Trp Tyr Leu Gly Ser Ser Ala Ala Asp Ala
    290   295   300
Asp Asn Thr Lys Phe Ala Asn Thr Ser Gly Met Ser Leu Leu Asp Phe

305 310 315 320
Arg Phe Asn Gln Glu Val Arg Asn Val Phe Arg Asn Asp Thr Ser Thr
325 330 335
Met Tyr Ala Leu Asp Ser Met Leu Thr Asn Thr Ala Ala Asp Tyr Thr
340 345 350
Gln Val Asn Asp Gln Val Thr Phe Ile Asp Asn His Asp Met Asp Arg
355 360 365
Phe Lys Thr Gly Thr Leu Asn Asn Arg Arg Leu Glu Gln Ala Leu Ala
370 375 380
Phe Thr Leu Thr Ser Arg Gly Val Pro Ala Ile Tyr Tyr Gly Thr Glu
385 390 395 400
Gln Tyr Met Thr Gly Asn Gly Asp Pro Asp Asn Arg Ala Met Met Thr
405 410 415
Ser Phe Ser Thr Ser Thr Thr Ala Phe Asn Val Ile Ser Lys Leu Ala
420 425 430
Pro Leu Arg Lys Ser Asn Pro Ala Ile Ala Tyr Gly Thr Thr Gln Gln
435 440 445
Arg Trp Ile Asn Asn Asp Val Tyr Ile Tyr Glu Arg Lys Phe Gly Asn
450 455 460
Asn Val Ala Val Val Ala Val Asn Arg Asn Leu Ser Thr Ala Thr Ser
465 470 475 480
Ile Ser Gly Leu Val Thr Ser Leu Pro Ser Gly Thr Tyr Ser Asp Val
485 490 495
Leu Gly Gly Thr Leu Ser Gly Asn Ser Ile Thr Ala Ser Ser Gly Asn
500 505 510
Val Ala Thr Phe Thr Leu Gly Ala Gly Ala Ala Ala Val Trp Gln Tyr
515 520 525
Thr Glu Thr Ser Ser Ala Thr Pro Val Leu Gly His Val Gly Pro Met
530 535 540
Met Gly Gln Pro Gly Asn Glu Val Thr Ile Asp Gly Arg Gly Phe Gly
545 550 555 560
Ser Ser Ala Gly Thr Val Tyr Phe Gly Thr Thr Ala Val Thr Gly Ser
565 570 575
Asn Ile Ile Ser Trp Glu Asp Thr Gln Ile Lys Val Lys Val Pro Ser
580 585 590
Val Ala Ala Gly Thr Tyr Ala Val Lys Val Thr Asn Ser Ala Gly Thr
595 600 605
Ser Asn Ser Tyr Ser Gly Phe Thr Ile Leu Thr Gly Thr Gln Val Ser
610 615 620
Val Arg Phe Ile Val Asn Asn Ala Thr Thr Ser Leu Gly Gln Asn Val
625 630 635 640
Tyr Leu Thr Gly Asn Val Ala Glu Leu Gly Asn Trp Thr Thr Gly Ser
645 650 655
Ser Ala Ile Gly Pro Leu Phe Asn Gln Ile Ile Lys Val Tyr Pro Thr
660 665 670
Trp Tyr Tyr Asp Val Ser Val Pro Ala Gly Thr Ala Leu Glu Phe Lys
675 680 685
Phe Phe Lys Lys Ser Gly Ser Thr Val Thr Trp Glu Ser Gly Ser Asn
690 695 700
His Thr Phe Thr Ala Pro Ala Ser Gly Thr Ala Thr Val Thr Val Asp
705 710 715 720
Trp Gln

<210> 426
<211> 2139
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 426
atgattagaa aattcacaat tgcgggtatc gttaacgaga atacgatcca gtacaaggta 60
acgcccgaaa cacttatctc cgatattgag gacataaaag tgtacgacgg cgatcgggag 120
attaaggtta atgaaatctc tactctcggt aagaaatccg ttatgggata cgtaacgctt 180
gacgagtcgc ttgatctttc gaagaattac agaatagagg tgaaaggata cggctcgaag 240
atcgctgttc ctaccgatat ttttgacagc gagtacttcg ctgaaacgta ccactatgac 300

```
ggagacgatc tcggagctgt gataaacgga gatacaacaa catttaaggt ttgggctccg     360
actgcgtctg aggttattct caatctcttt gaaaaaggcg acggagtaga ggcttacaag     420
agcgtttcga tgactcgcgg tgaaaagggc gtgtggtcac atacagagga atgcggtcat     480
ggaacgtact atacttatac tgtctcgaca tccgcaggtg ttcaggaagc tgttgaccca     540
tatgctaaat ccgcaggtgt gaacggaaac cgcggaatgg ttgttgacct ttccctgaca     600
aatcccgaag gatgggaaaa cgatcagctt gagaataaga tagagaagta ttcagatgca     660
ataatctggg aagtacacgt tcgtgacttc tcgaacaaga tcgaatcttc cgaatataag     720
ggcaagtacc ttgcgttcac ggagagaggt cttgttaacg agcacggtaa atctgtcggt     780
gttgattacc ttgtagacct tggtataaca cacgttcatc tgcttcctgt atatgactat     840
gctacagtag atgaatcctc gactgagccg cagtttaact ggggatacga cccgaagaat     900
tacaacgttc ccgagggcag ctattcaaca gatccgtaca acggcgaggt tcgtattcgc     960
gagttcaagg aaatggtaaa ggcactccac gaagcgggca ttggtgttat catggacgta    1020
gtatacaatc atactcatga cgcaaacagc tcgttcaata agatagttcc gtactactat    1080
tacagataca cctcctcagg cgctaactca tctgcaagcg gatgcggaaa cgataccgcc    1140
tccgagagat atatgtacgg caagttcatg gttgagtcca ctgcatattg ggtaagtgaa    1200
tacaaccttg acggacttcg cttcgacctt atgggactcc atgacgtaga aacgatgcag    1260
gaggttgaga gtgcagttca cgccatcaat ccgaatgcga tcatttacgg cgaaggatgg    1320
acaatgggcg cgactgttga cggaagcgct caggcaaatc agtcaaatat ctcaaagatt    1380
acgcctacag gtgacgcgat cggtgcagtc gctgtattca acgacgtgat cagagacggt    1440
ctcaagggaa gcgtatttga aaaaacaggc cgcggcttta tcaacggatc tgccaagaat    1500
acaataaaca aggttatgtt cggtatcaga ggcggtcagg gtgttggtca gggctggaca    1560
gttgaagacg gaatggtcat caactatatg agtgctcacg ataacaatac gctgtgggat    1620
aagctccttc tctcgaatcc taacgtgagc gacgagaaga gaaacgatat gaataacctt    1680
ggcgccgcaa tcatcatgat ctctcgcgga actccgttct ggcaggcggg agaggaaatg    1740
ctccgcacaa agggaggaga cgagaacagc tacaagtcca gtgatgaggt aaacaatatc    1800
aactggtccg tacttgctga cggtacgcgc gagtatgaaa caatgcagta ttacaagggt    1860
ctcatcgaga tgcgtaaggc atacggatta ttctcagata tgacaaccga gatcactcac    1920
accgactccg gcaacggaat ccttacaata agtatcaaag cacagaacgg tgaagaagca    1980
ctcgtaatta ttaatccgaa cgagtttgag tacacgcatg agcttgacgg cgagtggaaa    2040
ctcatatgtg acgctgaccg tgcaggtgca gaggttctcg aggctgagag cggaagcgtt    2100
aaggttggcg cgatcagtat taaggtttac gtaaagtga                           2139

<210> 427
<211> 712
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 427
Met Ile Arg Lys Phe Thr Ile Ala Gly Ile Val Asn Glu Asn Thr Ile
 1               5                  10                  15
Gln Tyr Lys Val Thr Pro Glu Thr Leu Ile Ser Asp Ile Glu Asp Ile
            20                  25                  30
Lys Val Tyr Asp Gly Asp Arg Glu Ile Lys Val Asn Glu Ile Ser Thr
        35                  40                  45
Leu Gly Lys Lys Ser Val Met Gly Tyr Val Thr Leu Asp Glu Ser Leu
    50                  55                  60
Asp Leu Ser Lys Asn Tyr Arg Ile Glu Val Lys Gly Tyr Gly Ser Lys
65                  70                  75                  80
Ile Ala Val Pro Thr Asp Ile Phe Asp Ser Glu Tyr Phe Ala Glu Thr
                85                  90                  95
Tyr His Tyr Asp Gly Asp Asp Leu Gly Ala Val Ile Asn Gly Asp Thr
            100                 105                 110
Thr Thr Phe Lys Val Trp Ala Pro Thr Ala Ser Glu Val Ile Leu Asn
        115                 120                 125
Leu Phe Glu Lys Gly Asp Gly Val Glu Ala Tyr Lys Ser Val Ser Met
    130                 135                 140
Thr Arg Gly Glu Lys Gly Val Trp Ser His Thr Glu Glu Cys Gly His
145                 150                 155                 160
Gly Thr Tyr Tyr Thr Tyr Thr Val Ser Thr Ser Ala Gly Val Gln Glu
                165                 170                 175
Ala Val Asp Pro Tyr Ala Lys Ser Ala Gly Val Asn Gly Asn Arg Gly
            180                 185                 190
Met Val Val Asp Leu Ser Leu Thr Asn Pro Glu Gly Trp Glu Asn Asp
        195                 200                 205
Gln Leu Glu Asn Lys Ile Glu Lys Tyr Ser Asp Ala Ile Ile Trp Glu
```

```
    210                 215                 220
Val His Val Arg Asp Phe Ser Asn Lys Ile Glu Ser Ser Glu Tyr Lys
225                 230                 235                 240
Gly Lys Tyr Leu Ala Phe Thr Glu Arg Gly Leu Val Asn Glu His Gly
                245                 250                 255
Lys Ser Val Gly Val Asp Tyr Leu Val Asp Leu Gly Ile Thr His Val
            260                 265                 270
His Leu Leu Pro Val Tyr Asp Tyr Ala Thr Val Asp Glu Ser Ser Thr
        275                 280                 285
Glu Pro Gln Phe Asn Trp Gly Tyr Asp Pro Lys Asn Tyr Asn Val Pro
    290                 295                 300
Glu Gly Ser Tyr Ser Thr Asp Pro Tyr Asn Gly Glu Val Arg Ile Arg
305                 310                 315                 320
Glu Phe Lys Glu Met Val Lys Ala Leu His Glu Ala Gly Ile Gly Val
                325                 330                 335
Ile Met Asp Val Val Tyr Asn His Thr His Asp Ala Asn Ser Ser Phe
            340                 345                 350
Asn Lys Ile Val Pro Tyr Tyr Tyr Tyr Arg Tyr Thr Ser Ser Gly Ala
        355                 360                 365
Asn Ser Ser Ala Ser Gly Cys Gly Asn Asp Thr Ala Ser Glu Arg Tyr
    370                 375                 380
Met Tyr Gly Lys Phe Met Val Glu Ser Thr Ala Tyr Trp Val Ser Glu
385                 390                 395                 400
Tyr Asn Leu Asp Gly Leu Arg Phe Asp Leu Met Gly Leu His Asp Val
                405                 410                 415
Glu Thr Met Gln Glu Val Glu Ser Ala Val His Ala Ile Asn Pro Asn
            420                 425                 430
Ala Ile Ile Tyr Gly Glu Gly Trp Thr Met Gly Ala Thr Val Asp Gly
        435                 440                 445
Ser Ala Gln Ala Asn Gln Ser Asn Ile Ser Lys Ile Thr Pro Thr Gly
    450                 455                 460
Asp Ala Ile Gly Ala Val Ala Val Phe Asn Asp Val Ile Arg Asp Gly
465                 470                 475                 480
Leu Lys Gly Ser Val Phe Glu Lys Thr Gly Arg Gly Phe Ile Asn Gly
                485                 490                 495
Ser Ala Lys Asn Thr Ile Asn Lys Val Met Phe Gly Ile Arg Gly Gly
            500                 505                 510
Gln Gly Val Gly Gln Gly Trp Thr Val Glu Asp Gly Met Val Ile Asn
        515                 520                 525
Tyr Met Ser Ala His Asp Asn Asn Thr Leu Trp Asp Lys Leu Leu Leu
    530                 535                 540
Ser Asn Pro Asn Val Ser Asp Glu Lys Arg Asn Asp Met Asn Asn Leu
545                 550                 555                 560
Gly Ala Ala Ile Ile Met Ile Ser Arg Gly Thr Pro Phe Trp Gln Ala
                565                 570                 575
Gly Glu Glu Met Leu Arg Thr Lys Gly Gly Asp Glu Asn Ser Tyr Lys
            580                 585                 590
Ser Ser Asp Glu Val Asn Asn Ile Asn Trp Ser Val Leu Ala Asp Gly
        595                 600                 605
Thr Arg Glu Tyr Glu Thr Met Gln Tyr Tyr Lys Gly Leu Ile Glu Met
    610                 615                 620
Arg Lys Ala Tyr Gly Leu Phe Ser Asp Met Thr Thr Glu Ile Thr His
625                 630                 635                 640
Thr Asp Ser Gly Asn Gly Ile Leu Thr Ile Ser Ile Lys Ala Gln Asn
                645                 650                 655
Gly Glu Glu Ala Leu Val Ile Ile Asn Pro Asn Glu Phe Glu Tyr Thr
            660                 665                 670
His Glu Leu Asp Gly Glu Trp Lys Leu Ile Cys Asp Ala Asp Arg Ala
        675                 680                 685
Gly Ala Glu Val Leu Glu Ala Glu Ser Gly Ser Val Lys Val Gly Ala
    690                 695                 700
Ile Ser Ile Lys Val Tyr Val Lys
705                 710

<210> 428
<211> 1425
<212> DNA
<213> Unknown
```

<220>
<223> Obtained from an environmental sample

<400> 428
atgcaagtaa tcagtaccat gctaaagcgt cagctattgt ggctggcatt cgttgttagt 60
tcagctgtgt tttcgtcggc tgctgttgcc gatgctattt tgcatgcctt tgactggcac 120
tatgatgaag ttgctgctaa agcgacagaa ataaaaaatc tgggttacaa ggcggtattg 180
gtagcgccgc cgttaaaatc gaacgcgggt aattgtgcct ggtggcagcg ttatcagccg 240
caggatatcc gtgtaattga tcattgcaaa ggtaacaaac aatcgtttgt aaatatgatc 300
aatgccctga acgatgccaa cccggcgcgt aaagttgatg tatacgccga tatcgtatta 360
aaccatatgg caaatgagcg taacggcgcg actgactttc ccggccaggc tgcagtaaat 420
agctatggca gtaacagcag ttactggaat aaccagcgcc tgtttggcaa tttaacccag 480
gggttgttcg gcgccagcga ctttaacccg gctaactgta tcagcaacta taacgatgta 540
tggcaggtgc agaactggcg gctgtgtggt ggtgcgggtg atgccggttt gccggatctc 600
aaccccaaca gctgggtggt gcagcaacag cgtaactact taactgcatt aaaagactta 660
ggcgtaaaag gctttagggt tgatgcggcc aagcatatga ccatttggca tattaatgaa 720
attttcacca gcagcattaa aaatggcatg tacgtgtttg gtgaaattat taccagcggt 780
ggagccggta ataacgaata cgatagcttc ttatcgcctt acctggctta taccgatcat 840
aaagcctacg attttccgct gtttagcgct attcgcagtg cctttggttt tggcggtagc 900
ttaagccaac tggttaaccc gttgtctaac ggccaggcgc tgcaaaacag ccgggcagta 960
acctttacta ttacgcatga tattcctacc aatgacggtt ttcgttatct gattatggat 1020
gctaccgacg agtatctggc ctatgcctac attatgggcc gtgatggtgg taaaccgctg 1080
atattcagcg acagcaccgg cactgataac aaccgctggg ttaatgccta taaagccgat 1140
catattagca aaatgctgaa tttccataac cgtatgcagg ggcagggcat ggaaatgtta 1200
gcctggaatg actgtgctat tttgttccgc cgcggccagg aaggcgtggt aggtattaat 1260
aaatgcagca gcaaccagag ctttaacatt aacaccaacg gccgctttta ctggtaccgt 1320
aattaccgcg atgtgttaag tggcggtaat ctggtgtata tcaatggcgg cagttacaac 1380
ttcagcattc cggcgcgtca ggccagaatg tggtatgccg actaa 1425

<210> 429
<211> 474
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(30)

<400> 429
Met Gln Val Ile Ser Thr Met Leu Lys Arg Gln Leu Leu Trp Leu Ala
 1   5   10   15
Phe Val Val Ser Ser Ala Val Phe Ser Ser Ala Ala Val Ala Asp Ala
            20   25   30
Ile Leu His Ala Phe Asp Trp His Tyr Asp Glu Val Ala Ala Lys Ala
        35   40   45
Thr Glu Ile Lys Asn Leu Gly Tyr Lys Ala Val Leu Val Ala Pro Pro
    50   55   60
Leu Lys Ser Asn Ala Gly Asn Cys Ala Trp Trp Gln Arg Tyr Gln Pro
65   70   75   80
Gln Asp Ile Arg Val Ile Asp His Cys Lys Gly Asn Lys Gln Ser Phe
                85   90   95
Val Asn Met Ile Asn Ala Leu Asn Asp Ala Asn Pro Ala Arg Lys Val
            100   105   110
Asp Val Tyr Ala Asp Ile Val Leu Asn His Met Ala Asn Glu Arg Asn
        115   120   125
Gly Ala Thr Asp Phe Pro Gly Gln Ala Ala Val Asn Ser Tyr Gly Ser
    130   135   140
Asn Ser Ser Tyr Trp Asn Asn Gln Arg Leu Phe Gly Asn Leu Thr Gln
145   150   155   160
Gly Leu Phe Gly Ala Ser Asp Phe Asn Pro Ala Asn Cys Ile Ser Asn
                165   170   175
Tyr Asn Asp Val Trp Gln Val Gln Asn Trp Arg Leu Cys Gly Gly Ala
            180   185   190
Gly Asp Ala Gly Leu Pro Asp Leu Asn Pro Asn Ser Trp Val Val Gln
        195   200   205

```
Gln Gln Arg Asn Tyr Leu Thr Ala Leu Lys Asp Leu Gly Val Lys Gly
    210                 215                 220
Phe Arg Val Asp Ala Ala Lys His Met Thr Ile Trp His Ile Asn Glu
225                 230                 235                 240
Ile Phe Thr Ser Ser Ile Lys Asn Gly Met Tyr Val Phe Gly Glu Ile
                245                 250                 255
Ile Thr Ser Gly Gly Ala Gly Asn Asn Glu Tyr Asp Ser Phe Leu Ser
            260                 265                 270
Pro Tyr Leu Ala Tyr Thr Asp His Lys Ala Tyr Asp Phe Pro Leu Phe
        275                 280                 285
Ser Ala Ile Arg Ser Ala Phe Gly Phe Gly Gly Ser Leu Ser Gln Leu
    290                 295                 300
Val Asn Pro Leu Ser Asn Gly Gln Ala Leu Gln Asn Ser Arg Ala Val
305                 310                 315                 320
Thr Phe Thr Ile Thr His Asp Ile Pro Thr Asn Asp Gly Phe Arg Tyr
                325                 330                 335
Leu Ile Met Asp Ala Thr Asp Glu Tyr Leu Ala Tyr Ala Tyr Ile Met
            340                 345                 350
Gly Arg Asp Gly Gly Lys Pro Leu Ile Phe Ser Asp Ser Thr Gly Thr
        355                 360                 365
Asp Asn Asn Arg Trp Val Asn Ala Tyr Lys Ala Asp His Ile Ser Lys
    370                 375                 380
Met Leu Asn Phe His Asn Arg Met Gln Gly Gln Gly Met Glu Met Leu
385                 390                 395                 400
Ala Trp Asn Asp Cys Ala Ile Leu Phe Arg Arg Gly Gln Glu Gly Val
                405                 410                 415
Val Gly Ile Asn Lys Cys Ser Ser Asn Gln Ser Phe Asn Ile Asn Thr
            420                 425                 430
Asn Gly Arg Phe Tyr Trp Tyr Arg Asn Tyr Arg Asp Val Leu Ser Gly
        435                 440                 445
Gly Asn Leu Val Tyr Ile Asn Gly Gly Ser Tyr Asn Phe Ser Ile Pro
    450                 455                 460
Ala Arg Gln Ala Arg Met Trp Tyr Ala Asp
465                 470
```

<210> 430
<211> 1650
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 430
atgtccctat tcaaaaaaag ctttccgtgg attttatccc tacttctttt gttttcgttt 60
attgctcctt tttccattca aacagaaaaa gtccgcgctg gaagtgtgcc ggtaaatggc 120
acaatgatgc aatatttcga atggtacctt ccagacgatg gaacactatg gacgaaagta 180
gcaaataacg ctcaatcttt agcgaatctt ggcattactg ccctttggct tccccctgcc 240
tataaaggaa caagcagcag tgacgttgga tatggcgttt atgatttata tgaccttgga 300
gagtttaatc aaaaaggaac tgtccgaaca aaatacggga caaaaacaca atatatccaa 360
gcaatccaag cggcgcatac agcagggatg cgagtatatg cagatgtcgt ctttaaccat 420
aaagccggtg cagatggaac agaactagtc gatgcagtag aagtaaatcc ttctgaccgc 480
aatcaagaaa tatcaggaac atatcaaatc caagcgtgga caaaatttga ttttcctggt 540
cgtggaaaca cctattctag ttttaaatgg cgttggtatc atttcgatgg aacggactgg 600
gatgagagta gaaaactaaa tcgtatttac aagttccgcg gcacgggaaa agcatgggat 660
tgggaagtag atacagaaaa cgggaattat gactatctca tgtatgcaga tttagatatg 720
gatcatccag aggttgtatc cgaactaaaa aattggggaa agtggtatgt aaccacaacc 780
aatatcgacg gattccgtct ggatgcagtg aagcatatta aatatagctt tttcccggac 840
tggctatcgt acgtacgaac ccaaacacaa aagcctcttt ttgccgttgg ggaattttgg 900
agctatgaca ttagcaagtt gcacaactat attacaaaga cgaacggctc tatgtcccta 960
ttcgatgccc cgctgcataa caatttttat atagcatcga aatcaagcgg ttattttgat 1020
atgcgcacat tactcaacaa cacattgatg aaagatcagc ctacattagc agtcacatta 1080
gtggataatc acgatactga gccagggcaa tctctgcagt catgggtcga gccatggttt 1140
aaaccgttag cttacgcatt tatcttgacc cgccaagaag gttatccttg cgtcttttat 1200
ggagattact atggtattcc aaaatacaac attcctgcgc tgaaaagcaa acttgatccg 1260
ctgttaattg ccagaagaga ttatgcctat ggaacacagc acgactatat tgacagtgcg 1320
gatattatcg gttggacgcg ggaaggagtg gctgaaaaag caaattcagg actggctgca 1380
ctcattaccg acgggcctgg cggaagcaaa tggatgtatg ttggaaaaca acacgctggc 1440

aaaacgtttt atgatttaac cggcaatcga agtgatacag tgacaatcaa tgctgatgga 1500
tggggagaat ttaaagtcaa tggagggtct gtatccatat gggttccaaa aatatcaacc 1560
acttcccaaa taacatttac tgtaaataac gccacaaccg tttggggaca aaatgtatac 1620
gttgtcggga atatttcgca gctggggaac 1650

<210> 431
<211> 550
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(33)

<400> 431
Met Ser Leu Phe Lys Lys Ser Phe Pro Trp Ile Leu Ser Leu Leu Leu
 1   5   10   15
Leu Phe Ser Phe Ile Ala Pro Phe Ser Ile Gln Thr Glu Lys Val Arg
            20   25   30
Ala Gly Ser Val Pro Val Asn Gly Thr Met Met Gln Tyr Phe Glu Trp
        35   40   45
Tyr Leu Pro Asp Asp Gly Thr Leu Trp Thr Lys Val Ala Asn Asn Ala
    50   55   60
Gln Ser Leu Ala Asn Leu Gly Ile Thr Ala Leu Trp Leu Pro Pro Ala
65   70   75   80
Tyr Lys Gly Thr Ser Ser Ser Asp Val Gly Tyr Gly Val Tyr Asp Leu
                85   90   95
Tyr Asp Leu Gly Glu Phe Asn Gln Lys Gly Thr Val Arg Thr Lys Tyr
            100   105   110
Gly Thr Lys Thr Gln Tyr Ile Gln Ala Ile Gln Ala Ala His Thr Ala
        115   120   125
Gly Met Arg Val Tyr Ala Asp Val Val Phe Asn His Lys Ala Gly Ala
    130   135   140
Asp Gly Thr Glu Leu Val Asp Ala Val Glu Val Asn Pro Ser Asp Arg
145   150   155   160
Asn Gln Glu Ile Ser Gly Thr Tyr Gln Ile Gln Ala Trp Thr Lys Phe
                165   170   175
Asp Phe Pro Gly Arg Gly Asn Thr Tyr Ser Ser Phe Lys Trp Arg Trp
            180   185   190
Tyr His Phe Asp Gly Thr Asp Trp Asp Glu Ser Arg Lys Leu Asn Arg
        195   200   205
Ile Tyr Lys Phe Arg Gly Thr Gly Lys Ala Trp Asp Trp Glu Val Asp
    210   215   220
Thr Glu Asn Gly Asn Tyr Asp Tyr Leu Met Tyr Ala Asp Leu Asp Met
225   230   235   240
Asp His Pro Glu Val Val Ser Glu Leu Lys Asn Trp Gly Lys Trp Tyr
                245   250   255
Val Thr Thr Thr Asn Ile Asp Gly Phe Arg Leu Asp Ala Val Lys His
            260   265   270
Ile Lys Tyr Ser Phe Phe Pro Asp Trp Leu Ser Tyr Val Arg Thr Gln
        275   280   285
Thr Gln Lys Pro Leu Phe Ala Val Gly Glu Phe Trp Ser Tyr Asp Ile
    290   295   300
Ser Lys Leu His Asn Tyr Ile Thr Lys Thr Asn Gly Ser Met Ser Leu
305   310   315   320
Phe Asp Ala Pro Leu His Asn Asn Phe Tyr Ile Ala Ser Lys Ser Ser
                325   330   335
Gly Tyr Phe Asp Met Arg Thr Leu Leu Asn Asn Thr Leu Met Lys Asp
            340   345   350
Gln Pro Thr Leu Ala Val Thr Leu Val Asp Asn His Asp Thr Glu Pro
        355   360   365
Gly Gln Ser Leu Gln Ser Trp Val Glu Pro Trp Phe Lys Pro Leu Ala
    370   375   380
Tyr Ala Phe Ile Leu Thr Arg Gln Glu Gly Tyr Pro Cys Val Phe Tyr
385   390   395   400
Gly Asp Tyr Tyr Gly Ile Pro Lys Tyr Asn Ile Pro Ala Leu Lys Ser

```
                405                 410                 415
Lys Leu Asp Pro Leu Leu Ile Ala Arg Arg Asp Tyr Ala Tyr Gly Thr
            420                 425                 430
Gln His Asp Tyr Ile Asp Ser Ala Asp Ile Ile Gly Trp Thr Arg Glu
        435                 440                 445
Gly Val Ala Glu Lys Ala Asn Ser Gly Leu Ala Ala Leu Ile Thr Asp
    450                 455                 460
Gly Pro Gly Gly Ser Lys Trp Met Tyr Val Gly Lys Gln His Ala Gly
465                 470                 475                 480
Lys Thr Phe Tyr Asp Leu Thr Gly Asn Arg Ser Asp Thr Val Thr Ile
                485                 490                 495
Asn Ala Asp Gly Trp Gly Glu Phe Lys Val Asn Gly Gly Ser Val Ser
            500                 505                 510
Ile Trp Val Pro Lys Ile Ser Thr Thr Ser Gln Ile Thr Phe Thr Val
        515                 520                 525
Asn Asn Ala Thr Thr Val Trp Gly Gln Asn Val Tyr Val Val Gly Asn
    530                 535                 540
Ile Ser Gln Leu Gly Asn
545                 550
```

<210> 432
<211> 1752
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 432

```
atgtccgaaa gaggcgtcag gagagccgtg cgcaccgccc tggtcgggtt ggcggcggcc      60
gccaccgccg cggtgacgct cggcgctccc accgcccagg ccgcgccggc cgggaacaag     120
gacgtcaccg ccgtcctgtt ctcctggaac ttcaactcca tcgcccgcga gtgcaccgag     180
cgactcggcc ccgcgggcta cggcttcgtg caggtctcgc cgccccagga gcacatccag     240
ggctccgcgt ggtggaccca gtaccagccc gtgagttaca acatctccgg ccggctcggc     300
aacgagcagc agttccgctc gatggtcaac acctgtaaca acgcgggtgt cggcgtcatc     360
gtggacgcgg tcatcaacca catgtccgcc ggctcgggca ccggtaccgg cggcacctcc     420
tactccaagt acaactaccc cggcctctac agcggcaacg acttcggttc ctgccgcgag     480
gacatccggg gcggtgacta caccaacgac cgctggcggg tccagaactg cgagctggtc     540
ggcctcgccg acctggacac cggcagcgcc tacgtccagc agcagatcgc cggctacatg     600
aaccgtctgc tgggctgggg cgtcgccggt ttccgggtgg acgccgtcaa gcacatcccg     660
gcccggcaca tggagcagat ccgcgcccgg gtcaacggcg gcgacgtctt ctggaaccag     720
gaggtcatct acggcgccgg cgaggcgatc acgcccgagg agtacctcaa caccggtcac     780
gtgcaggagt tccgctacgc gttcgacctc aagcggatgt tccagggcga ccggatcgcc     840
aatctgcaga acttcggcga gtcatggggc tacatgccct ccaaccgctc cggcgtcttc     900
gtcgacaacc acgacaccga gcgcaacggc tccaccctca gctacaagga caacgccgcc     960
tacaccctgg ccaacgtctt catgctggcc tggccctacg gcagcccgga cgtgcactcc    1020
ggctacgagt tcaccaactt cgacgccggt ccgcccaaca acggcaacgt gaccgcctgc    1080
caccagagcg ggtggaagtg ccagcacgcc tggcaggaga tctcctccat ggtgggcttc    1140
cgcaacgcca cccgcggcca ggccgtcacc aactggtggt ccaacggcaa caacgccatc    1200
gccttcggcc ggggcaaccg cggctacgtg gcgatcaacc acgagaacac cacactgaac    1260
cggaccttcc agacctccct gccggccggt gactactgca acgtgcagaa cggcaccacc    1320
gtcaccgtca acggcgccgg gcagttcacc gcgagcctcg gcgcccgcac cgccctcgcc    1380
ctgcacgtcg acgcccgcaa ttgcgccggc ggcggcaccg ggggcaacgg cggcaccggg    1440
ggaggcgaga acccggtgac cggcggtgcc tccttcggcg tcgacgcgac cacgcagatg    1500
ggccagaaca tccacgtcgc gggcaacatc cccgcgctgg gcgactggaa caccgccaac    1560
gccccccgca tgagcgccga cacctacccg gtctggcggc tggaactgaa cctgccgccc    1620
ggcaccacct tccagtacaa gtacatccgc agggacgcca acggcaacgt gacctgggaa    1680
tccgggaaca accgcaccgc caccgtgccc tcctccggcc gggtgacgct caacgacacc    1740
tggcgcaact ga                                                        1752
```

<210> 433
<211> 583
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(34)

<400> 433

```
Met Ser Glu Arg Gly Val Arg Arg Ala Val Arg Thr Ala Leu Val Gly
 1               5                  10                  15
Leu Ala Ala Ala Ala Thr Ala Ala Val Thr Leu Gly Ala Pro Thr Ala
            20                  25                  30
Gln Ala Ala Pro Ala Gly Asn Lys Asp Val Thr Ala Val Leu Phe Ser
        35                  40                  45
Trp Asn Phe Asn Ser Ile Ala Arg Glu Cys Thr Glu Arg Leu Gly Pro
    50                  55                  60
Ala Gly Tyr Gly Phe Val Gln Val Ser Pro Pro Gln Glu His Ile Gln
65                  70                  75                  80
Gly Ser Ala Trp Trp Thr Gln Tyr Gln Pro Val Ser Tyr Asn Ile Ser
                85                  90                  95
Gly Arg Leu Gly Asn Glu Gln Gln Phe Arg Ser Met Val Asn Thr Cys
            100                 105                 110
Asn Asn Ala Gly Val Gly Val Ile Val Asp Ala Val Ile Asn His Met
        115                 120                 125
Ser Ala Gly Ser Gly Thr Gly Thr Gly Gly Thr Ser Tyr Ser Lys Tyr
    130                 135                 140
Asn Tyr Pro Gly Leu Tyr Ser Gly Asn Asp Phe Gly Ser Cys Arg Glu
145                 150                 155                 160
Asp Ile Arg Gly Gly Asp Tyr Thr Asn Asp Arg Trp Arg Val Gln Asn
                165                 170                 175
Cys Glu Leu Val Gly Leu Ala Asp Leu Asp Thr Gly Ser Ala Tyr Val
            180                 185                 190
Gln Gln Gln Ile Ala Gly Tyr Met Asn Arg Leu Leu Gly Trp Gly Val
        195                 200                 205
Ala Gly Phe Arg Val Asp Ala Val Lys His Ile Pro Ala Arg His Met
    210                 215                 220
Glu Gln Ile Arg Ala Arg Val Asn Gly Gly Asp Val Phe Trp Asn Gln
225                 230                 235                 240
Glu Val Ile Tyr Gly Ala Gly Glu Ala Ile Thr Pro Glu Glu Tyr Leu
                245                 250                 255
Asn Thr Gly His Val Gln Glu Phe Arg Tyr Ala Phe Asp Leu Lys Arg
            260                 265                 270
Met Phe Gln Gly Asp Arg Ile Ala Asn Leu Gln Asn Phe Gly Glu Ser
        275                 280                 285
Trp Gly Tyr Met Pro Ser Asn Arg Ser Gly Val Phe Val Asp Asn His
    290                 295                 300
Asp Thr Glu Arg Asn Gly Ser Thr Leu Ser Tyr Lys Asp Asn Ala Ala
305                 310                 315                 320
Tyr Thr Leu Ala Asn Val Phe Met Leu Ala Trp Pro Tyr Gly Ser Pro
                325                 330                 335
Asp Val His Ser Gly Tyr Glu Phe Thr Asn Phe Asp Ala Gly Pro Pro
            340                 345                 350
Asn Asn Gly Asn Val Thr Ala Cys His Gln Ser Gly Trp Lys Cys Gln
        355                 360                 365
His Ala Trp Gln Glu Ile Ser Ser Met Val Gly Phe Arg Asn Ala Thr
    370                 375                 380
Arg Gly Gln Ala Val Thr Asn Trp Trp Ser Asn Gly Asn Asn Ala Ile
385                 390                 395                 400
Ala Phe Gly Arg Gly Asn Arg Gly Tyr Val Ala Ile Asn His Glu Asn
                405                 410                 415
Thr Thr Leu Asn Arg Thr Phe Gln Thr Ser Leu Pro Ala Gly Asp Tyr
            420                 425                 430
Cys Asn Val Gln Asn Gly Thr Thr Val Thr Val Asn Gly Ala Gly Gln
        435                 440                 445
Phe Thr Ala Ser Leu Gly Ala Arg Thr Ala Leu Ala Leu His Val Asp
    450                 455                 460
Ala Arg Asn Cys Ala Gly Gly Gly Thr Gly Gly Asn Gly Gly Thr Gly
465                 470                 475                 480
Gly Gly Glu Asn Pro Val Thr Gly Gly Ala Ser Phe Gly Val Asp Ala
                485                 490                 495
Thr Thr Gln Met Gly Gln Asn Ile His Val Ala Gly Asn Ile Pro Ala
            500                 505                 510
```

```
Leu Gly Asp Trp Asn Thr Ala Asn Ala Pro Arg Met Ser Ala Asp Thr
        515                 520                 525
Tyr Pro Val Trp Arg Leu Glu Leu Asn Leu Pro Pro Gly Thr Thr Phe
    530                 535                 540
Gln Tyr Lys Tyr Ile Arg Arg Asp Ala Asn Gly Asn Val Thr Trp Glu
545                 550                 555                 560
Ser Gly Asn Asn Arg Thr Ala Thr Val Pro Ser Ser Gly Arg Val Thr
                565                 570                 575
Leu Asn Asp Thr Trp Arg Asn
            580
```

<210> 434
<211> 1302
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 434
atgagcaata tgtataaacg caactcaatg tttttattta tagttctttt tgttttgtta 60
tcttttgtgc cagcagtggc agacccaccc caggaagtat cagaaggcgt tatgcttcag 120
ggttttggct ggaacagtca agttagaggt tcgccctcta aatggtacac tcttattgaa 180
agtaaagcag acagcatcaa aaatctaggc gttgattttg tgtggttccc gcctgtagct 240
cgttctgtaa gccctcaggg atacctccct ggtgactact atgatttagg cagtgacgat 300
taccctacat tttatggcac caaagatcag ctaaaggcag ctttaaaagc tctaaacgac 360
agaaacattg cacctgttgc cgacattgta gttaatcata gatgtgctgg caagcaagat 420
gagaacggca tctggaatat attccactat ccttcgggca aggcacaatg ggaacaatgg 480
gcggtctgtg aaggtcagtt tggtggaaca ggcaatcctg acacaggaga taattttcat 540
gctgcacctg acattgacca cacaaacaag caggtccaac aagacataag cgattggatg 600
aactggctaa aagagctggg ctttagagct tggcgctacg actattctag aggttatgct 660
gctaaatatg cgggtcttta tgacaaaaac acacaacctt tgttctctgt tggcgaactt 720
tggacgaata tggctttcga aggctcctgg ctcaagccaa atcaagatgc tcacaggcaa 780
ttacttgttg actggcttga cgagaatcct agtccagttg ccacaatatt tgactttacc 840
accaaaggga ttttacaggt agccgtacac ggagaatacc aaagacttat tgactcaaac 900
aataatgctc caggccttat tggatggtgg cctcaaagag cagtaacatt tattgacaat 960
cacgacactg gctctcagca aagtcattgg ccatttgcag acgacaaagt tatgcagggt 1020
tatgcataca tcttaacaca ccccgggatt ccctgcctat tctgggagca cgtttacgac 1080
tggaatttgt atgatcaaat caagcaacta gttgatattc gcaagcaaaa taaaataaac 1140
agtacaagct cgctaaaaat cttaagagcc gagacgggcc tctacgcagc aattattgat 1200
gacaaggttg ttttaaaact tggttggcaa gattggtatc ctggcgaagg ctttgaacta 1260
aaagccagcg gagagcagta tgctgtgtgg gttaaaaact ag 1302

<210> 435
<211> 433
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(27)

<400> 435
```
Met Ser Asn Met Tyr Lys Arg Asn Ser Met Phe Leu Phe Ile Val Leu
 1               5                  10                  15
Phe Val Leu Leu Ser Phe Val Pro Ala Val Ala Asp Pro Pro Gln Glu
            20                  25                  30
Val Ser Glu Gly Val Met Leu Gln Gly Phe Gly Trp Asn Ser Gln Val
        35                  40                  45
Arg Gly Ser Pro Ser Lys Trp Tyr Thr Leu Ile Glu Ser Lys Ala Asp
    50                  55                  60
Ser Ile Lys Asn Leu Gly Val Asp Phe Val Trp Phe Pro Pro Val Ala
65                  70                  75                  80
Arg Ser Val Ser Pro Gln Gly Tyr Leu Pro Gly Asp Tyr Tyr Asp Leu
                85                  90                  95
Gly Ser Asp Asp Tyr Pro Thr Phe Tyr Gly Thr Lys Asp Gln Leu Lys
```

```
            100                 105                 110
Ala Ala Leu Lys Ala Leu Asn Asp Arg Asn Ile Ala Pro Val Ala Asp
        115                 120                 125
Ile Val Val Asn His Arg Cys Ala Gly Lys Gln Asp Glu Asn Gly Ile
    130                 135                 140
Trp Asn Ile Phe His Tyr Pro Ser Gly Lys Ala Gln Trp Glu Gln Trp
145                 150                 155                 160
Ala Val Cys Glu Gly Gln Phe Gly Gly Thr Gly Asn Pro Asp Thr Gly
                165                 170                 175
Asp Asn Phe His Ala Ala Pro Asp Ile Asp His Thr Asn Lys Gln Val
            180                 185                 190
Gln Gln Asp Ile Ser Asp Trp Met Asn Trp Leu Lys Glu Leu Gly Phe
        195                 200                 205
Arg Ala Trp Arg Tyr Asp Tyr Ser Arg Gly Tyr Ala Ala Lys Tyr Ala
    210                 215                 220
Gly Leu Tyr Asp Lys Asn Thr Gln Pro Leu Phe Ser Val Gly Glu Leu
225                 230                 235                 240
Trp Thr Asn Met Ala Phe Glu Gly Ser Trp Leu Lys Pro Asn Gln Asp
                245                 250                 255
Ala His Arg Gln Leu Leu Val Asp Trp Leu Asp Glu Asn Pro Ser Pro
            260                 265                 270
Val Ala Thr Ile Phe Asp Phe Thr Thr Lys Gly Ile Leu Gln Val Ala
        275                 280                 285
Val His Gly Glu Tyr Gln Arg Leu Ile Asp Ser Asn Asn Asn Ala Pro
    290                 295                 300
Gly Leu Ile Gly Trp Trp Pro Gln Arg Ala Val Thr Phe Ile Asp Asn
305                 310                 315                 320
His Asp Thr Gly Ser Gln Gln Ser His Trp Pro Phe Ala Asp Asp Lys
                325                 330                 335
Val Met Gln Gly Tyr Ala Tyr Ile Leu Thr His Pro Gly Ile Pro Cys
            340                 345                 350
Leu Phe Trp Glu His Val Tyr Asp Trp Asn Leu Tyr Asp Gln Ile Lys
        355                 360                 365
Gln Leu Val Asp Ile Arg Lys Gln Asn Lys Ile Asn Ser Thr Ser Ser
    370                 375                 380
Leu Lys Ile Leu Arg Ala Glu Thr Gly Leu Tyr Ala Ala Ile Ile Asp
385                 390                 395                 400
Asp Lys Val Val Leu Lys Leu Gly Trp Gln Asp Trp Tyr Pro Gly Glu
                405                 410                 415
Gly Phe Glu Leu Lys Ala Ser Gly Glu Gln Tyr Ala Val Trp Val Lys
            420                 425                 430
Asn
```

<210> 436
<211> 1311
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetically generated

<400> 436

```
atggccaagt actccgagct ggaaaagggc ggggtcataa tgcaggcgtt ctactgggac      60
gtgccttcag gaggaatatg gtgggacaca atacggcaga agataccgga gtggtacgat     120
gccggaatct ccgcaatatg gattcccccg gcgagcaagg gcatgggcgg cgcctattcg     180
atgggctacg acccctacga cttctttgac ctcggtgagt acgaccagaa gggaacggta     240
gagacgcgct ttggctccaa gcaggagctc gtgaacatga taaacaccgc ccacgcctat     300
ggcatgaagg taatagccga tatagtcatc aaccaccgcg ccggcggtga cctggagtgg     360
aacccttcg tgaacgacta tacctggacc gacttctcaa aggtcgcgtc gggtaaatac     420
acggccaact acctcgactt ccacccgaac gagctccatg cgggcgattc cggaacattt     480
ggaggctatc ccgacatatg ccacgacaag agctgggacc agtactggct ctgggccagc     540
caggagagct acgcggcata tctcaggagc atcggcatcg atgcctggcg cttcgactac     600
gtcaagggct atgctccctg ggtcgtcaag gactggctga actggtgggg aggctgggcg     660
gttggagagt actgggacac caacgtcgac gctgttctca actgggcata ctcgagcggt     720
gccaaggtct ttgacttcgc cctctactac aagatggatg aggcctttga caacaaaaac     780
attccagcgc tcgtctctgc ccttcagaac ggccagactg ttgtctcccg cgacccgttc     840
aaggccgtaa cctttgtagc aaaccacgac accgatataa tctggaacaa gtatccagcc     900
```

```
tacgcgttca tcctcaccta cgagggccag ccgacaatat tctaccgcga ctacgaggag     960
tggctcaaca aggataagct caagaacctc atctggatac atgagaacct cgccggagga    1020
agcaccgaca tagtctacta cgataacgat gaactcatct tcgtcaggaa cggctacggg    1080
gacaagccgg ggcttataac ctacatcaac ctaggctcga gcaaggccgg aaggtgggtt    1140
tatgtgccga agttcgcggg cgcgtgcatc cacgagtata ctggtaacct cggaggctgg    1200
gtagacaagt acgtctactc aagcggctgg gtctatctcg aagctccagc ttacgaccct    1260
gccaacgggc agtatggcta ctccgtgtgg agctactgcg gggtgggctg a             1311

<210> 437
<211> 436
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetically generated

<400> 437
Met Ala Lys Tyr Ser Glu Leu Glu Lys Gly Gly Val Ile Met Gln Ala
 1               5                  10                  15
Phe Tyr Trp Asp Val Pro Ser Gly Gly Ile Trp Trp Asp Thr Ile Arg
            20                  25                  30
Gln Lys Ile Pro Glu Trp Tyr Asp Ala Gly Ile Ser Ala Ile Trp Ile
        35                  40                  45
Pro Pro Ala Ser Lys Gly Met Gly Gly Ala Tyr Ser Met Gly Tyr Asp
    50                  55                  60
Pro Tyr Asp Phe Phe Asp Leu Gly Glu Tyr Asp Gln Lys Gly Thr Val
65                  70                  75                  80
Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Val Asn Met Ile Asn Thr
                85                  90                  95
Ala His Ala Tyr Gly Met Lys Val Ile Ala Asp Ile Val Ile Asn His
            100                 105                 110
Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Asn Asp Tyr Thr
        115                 120                 125
Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr
    130                 135                 140
Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe
145                 150                 155                 160
Gly Gly Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp
                165                 170                 175
Leu Trp Ala Ser Gln Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly
            180                 185                 190
Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Ala Pro Trp Val
        195                 200                 205
Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr
    210                 215                 220
Trp Asp Thr Asn Val Asp Ala Val Leu Asn Trp Ala Tyr Ser Ser Gly
225                 230                 235                 240
Ala Lys Val Phe Asp Phe Ala Leu Tyr Tyr Lys Met Asp Glu Ala Phe
                245                 250                 255
Asp Asn Lys Asn Ile Pro Ala Leu Val Ser Ala Leu Gln Asn Gly Gln
            260                 265                 270
Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn
        275                 280                 285
His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile
    290                 295                 300
Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu
305                 310                 315                 320
Trp Leu Asn Lys Asp Lys Leu Lys Asn Leu Ile Trp Ile His Glu Asn
                325                 330                 335
Leu Ala Gly Gly Ser Thr Asp Ile Val Tyr Tyr Asp Asn Asp Glu Leu
            340                 345                 350
Ile Phe Val Arg Asn Gly Tyr Gly Asp Lys Pro Gly Leu Ile Thr Tyr
        355                 360                 365
Ile Asn Leu Gly Ser Ser Lys Ala Gly Arg Trp Val Tyr Val Pro Lys
    370                 375                 380
Phe Ala Gly Ala Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp
385                 390                 395                 400
Val Asp Lys Tyr Val Tyr Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro
```

405 410 415
Ala Tyr Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr
420 425 430
Cys Gly Val Gly
435

<210> 438
<211> 1539
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 438
atgaaattta agcgcgctgc gggaatcatc cttcatccta ccagcttgcc ggggccggat 60
ggaattggtg accttggacc ggaagcatac cgctggatca actttctagc cgaatccggc 120
tgcaaactat ggcaggtgct gccgcttggc ccaaccggtt acggcgactc gccctatcag 180
tgtttctcag catttgccgg taatccctat ctggtcagcc cggcgctgtt gctggaaaac 240
ggtttgttac ggcggagtga tctgaccgac cgccctgaat ttaatccaat tcgaattgat 300
tacggtgaag ccatcacttg gaaattaaaa atactcgatc gcgcttacga aaactttaag 360
cgcctcaata atcctgaact aaaatatgaa tatcagcaat tcacggaaaa tgagacggac 420
tggttgaatg attttgccct ttttatggcc attaaggagg ctaacggggg ggtatcctgg 480
gacaactggc caaaggaact gcgtagccgc cagccggacg caatcgaaaa gttcaaacaa 540
accgaagccg accgcattca acgccatgcc tttcgacagt ttttattctt ccgccagtgg 600
cttgacctga aagcctacgc caatcaaaag aacatccaaa tcatcggtga tataccgatt 660
tttgtcgctt acgatagcgc agatgcttgg tcgaacccag agttgtttta tctgaatgaa 720
gaaggcaaac caactgttgt ggcaggcgtt ccaccggatt atttttcccc aacgggtcaa 780
ttatggggta acccgcttta caagtgggaa gttcatcgtc agcagaattt ccgctggtgg 840
attcgccgga tgcaagccac attacggctg gtagatattg tccgcctcga ccattttcgc 900
ggatttgccg gttattggga agtacccttt gggatgccca ccgccgaaat tggccgctgg 960
gtaaaggggc ccggcaaaga attattcaat gccattcggg atgcccttgg cgatttaccc 1020
atcatcgctg aagatttggg ggtcatcacc cccgatgtga ttgaactgcg cgattcgttt 1080
aacttgcccg gtatgaaggt gttccagttt gcttttacaa cggacccgct cgatcccttc 1140
ctgccgcaca actacgaaac caattgtgta gcctacactg gtacccacga caacgacacc 1200
gccttgggtt ggtatcaaag cgcaccggaa aaagaacgtg actttattcg ccgatatctc 1260
gcccgttccg gcgaagatat atcatgggac atgatccgcg ctgtgtggtc ttcagtggcc 1320
gtttttgcgc ttgcgcccat gcaagatatt ttgagccttg gaaatgaagc ccgcatgaat 1380
ctaccgggaa gagccagcgg caactggtgt tggagattac accccgaggc tttcaacccg 1440
gtaatattaa aacggctcaa agaagtcaac tacttgtacg cgcgcaccga cattgagcgc 1500
accgtacaga ccgaaatacc cccgaactat cccttataa 1539

<210> 439
<211> 496
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> DOMAIN
<222> (11)...(493)
<223> 4-alpha-glucanotransferase

<400> 439
Met Lys Phe Lys Arg Ala Ala Gly Ile Ile Leu His Pro Thr Ser Leu
1 5 10 15
Pro Gly Pro Asp Gly Ile Gly Asp Leu Gly Pro Glu Ala Tyr Arg Trp
20 25 30
Ile Asn Phe Leu Ala Glu Ser Gly Cys Lys Leu Trp Gln Val Leu Pro
35 40 45
Leu Gly Pro Thr Gly Tyr Gly Asp Ser Pro Tyr Gln Cys Phe Ser Ala
50 55 60
Phe Ala Gly Asn Pro Tyr Leu Val Ser Pro Ala Leu Leu Leu Glu Asn
65 70 75 80
Gly Leu Leu Arg Arg Ser Asp Leu Thr Asp Arg Pro Glu Phe Asn Pro
85 90 95
Ile Arg Ile Asp Tyr Gly Glu Ala Ile Thr Trp Lys Leu Lys Ile Leu

```
            100                 105                 110
Asp Arg Ala Tyr Glu Asn Phe Lys Arg Leu Asn Asn Pro Glu Leu Lys
        115                 120                 125
Tyr Glu Tyr Gln Gln Phe Thr Glu Asn Glu Thr Asp Trp Leu Asn Asp
    130                 135                 140
Phe Ala Leu Phe Met Ala Ile Lys Glu Ala Asn Gly Gly Val Ser Trp
145                 150                 155                 160
Asp Asn Trp Pro Lys Glu Leu Arg Ser Arg Gln Pro Asp Ala Ile Glu
                165                 170                 175
Lys Phe Lys Gln Thr Glu Ala Asp Arg Ile Gln Arg His Ala Phe Arg
            180                 185                 190
Gln Phe Leu Phe Phe Arg Gln Trp Leu Asp Leu Lys Ala Tyr Ala Asn
        195                 200                 205
Gln Lys Asn Ile Gln Ile Ile Gly Asp Ile Pro Ile Phe Val Ala Tyr
    210                 215                 220
Asp Ser Ala Asp Ala Trp Ser Asn Pro Glu Leu Phe Tyr Leu Asn Glu
225                 230                 235                 240
Glu Gly Lys Pro Thr Val Val Ala Gly Val Pro Pro Asp Tyr Phe Ser
                245                 250                 255
Pro Thr Gly Gln Leu Trp Gly Asn Pro Leu Tyr Lys Trp Glu Val His
            260                 265                 270
Arg Gln Gln Asn Phe Arg Trp Trp Ile Arg Arg Met Gln Ala Thr Leu
        275                 280                 285
Arg Leu Val Asp Ile Val Arg Leu Asp His Phe Arg Gly Phe Ala Gly
    290                 295                 300
Tyr Trp Glu Val Pro Phe Gly Met Pro Thr Ala Glu Ile Gly Arg Trp
305                 310                 315                 320
Val Lys Gly Pro Gly Lys Glu Leu Phe Asn Ala Ile Arg Asp Ala Leu
                325                 330                 335
Gly Asp Leu Pro Ile Ile Ala Glu Asp Leu Gly Val Ile Thr Pro Asp
            340                 345                 350
Val Ile Glu Leu Arg Asp Ser Phe Asn Leu Pro Gly Met Lys Val Phe
        355                 360                 365
Gln Phe Ala Phe Thr Thr Asp Pro Leu Asp Pro Phe Leu Pro His Asn
    370                 375                 380
Tyr Glu Thr Asn Cys Val Ala Tyr Thr Gly Thr His Asp Asn Asp Thr
385                 390                 395                 400
Ala Leu Gly Trp Tyr Gln Ser Ala Pro Glu Lys Glu Arg Asp Phe Ile
                405                 410                 415
Arg Arg Tyr Leu Ala Arg Ser Gly Glu Asp Ile Ser Trp Asp Met Ile
            420                 425                 430
Arg Ala Val Trp Ser Ser Val Ala Val Phe Ala Leu Ala Pro Met Gln
        435                 440                 445
Asp Ile Leu Ser Leu Gly Asn Glu Ala Arg Met Asn Leu Pro Gly Arg
    450                 455                 460
Ala Ser Gly Asn Trp Cys Trp Arg Leu His Pro Glu Ala Phe Asn Pro
465                 470                 475                 480
Val Ile Leu Lys Arg Leu Lys Glu Val Asn Tyr Leu Tyr Ala Arg Thr
                485                 490                 495
```

<210> 440
<211> 3447
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 440
atgaatcaaa tagttaattt caaatcgcat ttctatagaa aaatagccct tttattttca 60
atcactttta tatgggcggc aggctcgctg tctgcacagg attttatgat gcagggctgg 120
tactgggatt atcctaaagt ggactgcaat cagtatcagg gcccgtctct tgcatcagag 180
atggctgctc ttgccgagga gcaaagtgag gccggtttca cgatgatgtg gcttccgcca 240
ttgtcgaaag caagttttgg aaactgcagc aatggatacg atccccggga tctgtatgat 300
tatggccagg ctaccggccg gaccgggctt ggcacaggta cagaagtaga aacctggatg 360
acatcactca gcgagaacga tatttatccg atagcggatg tggtttataa tcaccgggac 420
gggggtgaat gggaggataa ccctgcagtg cgcgattaca tcataaacta tccaaataac 480
ctatcctgcg acggcatgcc ggcaacaccc tacccggtaa acggcaaagc gaggtatgta 540

ctgcctttgg gcggagacag cggcaataac gaaggggatt actttttaa atttgcatct 600
gcttcgggcg aaaccggttt taatggggca gaatataaac tgtttttccg caccagggat 660
accgggttta ttgatgatcc gatttttgaa tcagaaaata atggcggcgc tgactgcgat 720
caggataatg atcaaatatt tttaggcaga gatatttttg catcacagga agtgggcgga 780
agctgtaata cggatgaatt ttatcttgaa ttaaattcgg gagactttga tgctgacgga 840
gatatgcttg agatttatat agaggaggtt ggtggcggtg gtaccggaat tgatcagcgt 900
atttatggaa tttggtcggc cagcagaagt gaagacatac atgatgattt aatcgtgcaa 960
acccggactg attttacaaa catgccaagc gggaaagggg caatgacatt ccgccatttc 1020
aagccaaacg gcataacacc aacctgtatg acgggtgata aagattttcc tttctttttc 1080
tttgatattg aacaggagta tgatggaaac gaaggcggcg agtcaacccg caaggttttc 1140
gacgactgga accaatggtt atgggatacc ctgggcatca gaggatttcg cctggatgcg 1200
gtaaaacatt tcccggcatg gtttgtggga gtactgctga atgatttgca tgccgcaggc 1260
aaaaatcccc ccatggtagt aggtgaacat tttacgctgg atgcctccgt tttgagagac 1320
tggatagaca gggtttacga tgcaatgacg ccgggggcag cagaagcgat ccaggtaaga 1380
gcctttgatt ttgaattgag agacaggctt aaacaagccg ccgataaccc aagttatgat 1440
acgcgaaata tatttaacag cgggcttgta gacggagtgg gtatgagcgg tgcgcactcg 1500
gtcacgtttc tgaataacca tgattacagg acgccgggtg agcatatcat cagaaatcaa 1560
atgctggcat atgcttatat ccttaccaat aaccgcattg gcctgcccag tatttttcac 1620
cccgattatt atggcatcga tatttatggc caggatcatc cgatcgaaaa acaacaggaa 1680
gcgatagatg aactaatgca aatacacagg gattttatta ccggagctga atttgttgat 1740
tacctgaatc gttttgatac gccatatgat aacgcgtatc agcaatccgg cccatggaat 1800
catttgctct atcaaattca gggtggcgtt ggtgagaagg atgtgatagt gctcataaat 1860
tttggtgatg aacaattgcg gtttaaccat actattaata cagcgaatgc tccacttggt 1920
acaacatttt ctttggtagc tggagaatcg aattttccta atccagaggt ggaggttagc 1980
ccaaataata ttgaaaactc tctttttttc gatataccgg cgaactctta tgcggtttat 2040
gtacaaagtg atgcagtgga actgacaggt gaagagggct tccgtatgct gtctgtgcct 2100
gttccaacaa cttattcggc gtttcttgat cctatttgga cacagggagc tgatgaaggt 2160
gccaacgtat ccggcggaga tccaaatgta tttacctggg ataatacatc agagaacggc 2220
agttccgaaa actggagcgg actcaacgat ttgacgggta aaatagaagc cggcaccggg 2280
ttcctggttt atgtatataa gaatgatgag tttggtgttg aaggtatctt tcccaaaact 2340
ctgagagtaa gcggccccgc gcatggcgat acatcccctg agatgaatca gaatgacggc 2400
ggattcaccc tgcttggaaa ccctttttcc actacgatcg actttaatca attggacaaa 2460
aatggcttaa ccgatgtggc ttatgtgtgg gacccgaatg caggttcctg gcggtcatgg 2520
gatgcagacg gaaatgccgg ggatttggcc aacgggttga ttgctccttt ccagggcttc 2580
tttgtacagt caggttcagg aagcccatcc gtcacattta atgatcaaac caaaagctcc 2640
ggtggcgtat tctacggaaa agaaaaagat gaacaacgca gcatcgtccg ccttgagctg 2700
agtggtgaag agcttcaaaa ctctgcctgg ctgcgcttta gtgaatccgg ttcatatgaa 2760
tcacaggttc gcggcgatgc actgcaactg gagccgctga gccaaacttt tgcgaagctc 2820
gccatcgaaa agggcggcac actgtttgat atcagccacc ttccggtaat ggatgaagag 2880
tacagtgttc cggtacacat cagcgcaact aacagcgggg tttatacact aactgcaacc 2940
gattttgaac tgcctggtaa ccttgagctg actttccacg atcacgcggc cggaacttcc 3000
gttccgctga atgcagattt cagttacacc attgagcagg cacaggcggc taaagctgct 3060
gaactgagcc cgctggagcg tgtgaagcag ggcccggttg tagcaaaagt aactgcggat 3120
tcaccgcagt acagcattac ggtcaatagt tcggcagata ttggtggtac ggatgagttg 3180
cctgaacgcg tagccctccg ccagaactat ccaaacccat ttaacccggt taccgtaatc 3240
acatacgagc tgcctgaaaa cagcaatgta accctggaag ttttcgacat ggccggccgt 3300
aaagtatcta cactggtaaa cgaaaaggta agcgcagggt cacacgatgt tgcgttcgat 3360
gcctctcacc tttcaagcgg tgtatacatc taccggctgc atgccggcga gcatatgctg 3420
acgcggaaac tgactttaat aaagtga 3447

<210> 441
<211> 1148
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(32)

<221> DOMAIN
<222> (34)...(569)
<223> Alpha amylase, catalytic domain

<400> 441
Met Asn Gln Ile Val Asn Phe Lys Ser His Phe Tyr Arg Lys Ile Ala
 1   5   10   15

```
Leu Leu Phe Ser Ile Thr Phe Ile Trp Ala Ala Gly Ser Leu Ser Ala
            20                  25                  30
Gln Asp Phe Met Met Gln Gly Trp Tyr Trp Asp Tyr Pro Lys Val Asp
        35                  40                  45
Cys Asn Gln Tyr Gln Gly Pro Ser Leu Ala Ser Glu Met Ala Ala Leu
    50                  55                  60
Ala Glu Glu Gln Ser Glu Ala Gly Phe Thr Met Met Trp Leu Pro Pro
65                  70                  75                  80
Leu Ser Lys Ala Ser Phe Gly Asn Cys Ser Asn Gly Tyr Asp Pro Arg
                85                  90                  95
Asp Leu Tyr Asp Tyr Gly Gln Ala Thr Gly Arg Thr Gly Leu Gly Thr
            100                 105                 110
Gly Thr Glu Val Glu Thr Trp Met Thr Ser Leu Ser Glu Asn Asp Ile
        115                 120                 125
Tyr Pro Ile Ala Asp Val Val Tyr Asn His Arg Asp Gly Gly Glu Trp
    130                 135                 140
Glu Asp Asn Pro Ala Val Arg Asp Tyr Ile Ile Asn Tyr Pro Asn Asn
145                 150                 155                 160
Leu Ser Cys Asp Gly Met Pro Ala Thr Pro Tyr Pro Val Asn Gly Lys
                165                 170                 175
Ala Arg Tyr Val Leu Pro Leu Gly Gly Asp Ser Gly Asn Asn Glu Gly
            180                 185                 190
Asp Tyr Phe Phe Lys Phe Ala Ser Ala Ser Gly Glu Thr Gly Phe Asn
        195                 200                 205
Gly Ala Glu Tyr Lys Leu Phe Phe Arg Thr Arg Asp Thr Gly Phe Ile
    210                 215                 220
Asp Asp Pro Ile Phe Glu Ser Glu Asn Asn Gly Gly Ala Asp Cys Asp
225                 230                 235                 240
Gln Asp Asn Asp Gln Ile Phe Leu Gly Arg Asp Ile Phe Ala Ser Gln
                245                 250                 255
Glu Val Gly Gly Ser Cys Asn Thr Asp Glu Phe Tyr Leu Glu Leu Asn
            260                 265                 270
Ser Gly Asp Phe Asp Ala Asp Gly Asp Met Leu Glu Ile Tyr Ile Glu
        275                 280                 285
Glu Val Gly Gly Gly Gly Thr Gly Ile Asp Gln Arg Ile Tyr Gly Ile
    290                 295                 300
Trp Ser Ala Ser Arg Ser Glu Asp Ile His Asp Asp Leu Ile Val Gln
305                 310                 315                 320
Thr Arg Thr Asp Phe Thr Asn Met Pro Ser Gly Lys Gly Ala Met Thr
                325                 330                 335
Phe Arg His Phe Lys Pro Asn Gly Ile Thr Pro Thr Cys Met Thr Gly
            340                 345                 350
Asp Lys Asp Phe Pro Phe Phe Phe Phe Asp Ile Glu Gln Glu Tyr Asp
        355                 360                 365
Gly Asn Glu Gly Gly Glu Ser Thr Arg Lys Val Phe Asp Asp Trp Asn
    370                 375                 380
Gln Trp Leu Trp Asp Thr Leu Gly Ile Arg Gly Phe Arg Leu Asp Ala
385                 390                 395                 400
Val Lys His Phe Pro Ala Trp Phe Val Gly Val Leu Leu Asn Asp Leu
                405                 410                 415
His Ala Ala Gly Lys Asn Pro Pro Met Val Val Gly Glu His Phe Thr
            420                 425                 430
Leu Asp Ala Ser Val Leu Arg Asp Trp Ile Asp Arg Val Tyr Asp Ala
        435                 440                 445
Met Thr Pro Gly Ala Ala Glu Ala Ile Gln Val Arg Ala Phe Asp Phe
    450                 455                 460
Glu Leu Arg Asp Arg Leu Lys Gln Ala Ala Asp Asn Pro Ser Tyr Asp
465                 470                 475                 480
Thr Arg Asn Ile Phe Asn Ser Gly Leu Val Asp Gly Val Gly Met Ser
                485                 490                 495
Gly Ala His Ser Val Thr Phe Leu Asn Asn His Asp Tyr Arg Thr Pro
            500                 505                 510
Gly Glu His Ile Ile Arg Asn Gln Met Leu Ala Tyr Ala Tyr Ile Leu
        515                 520                 525
Thr Asn Asn Arg Ile Gly Leu Pro Ser Ile Phe His Pro Asp Tyr Tyr
    530                 535                 540
Gly Ile Asp Ile Tyr Gly Gln Asp His Pro Ile Glu Lys Gln Gln Glu
545                 550                 555                 560
```

Ala Ile Asp Glu Leu Met Gln Ile His Arg Asp Phe Ile Thr Gly Ala
565 570 575
Glu Phe Val Asp Tyr Leu Asn Arg Phe Asp Thr Pro Tyr Asp Asn Ala
580 585 590
Tyr Gln Gln Ser Gly Pro Trp Asn His Leu Leu Tyr Gln Ile Gln Gly
595 600 605
Gly Val Gly Glu Lys Asp Val Ile Val Leu Ile Asn Phe Gly Asp Glu
610 615 620
Gln Leu Arg Phe Asn His Thr Ile Asn Thr Ala Asn Ala Pro Leu Gly
625 630 635 640
Thr Thr Phe Ser Leu Val Ala Gly Glu Ser Asn Phe Pro Asn Pro Glu
645 650 655
Val Glu Val Ser Pro Asn Asn Ile Glu Asn Ser Leu Phe Phe Asp Ile
660 665 670
Pro Ala Asn Ser Tyr Ala Val Tyr Val Gln Ser Asp Ala Val Glu Leu
675 680 685
Thr Gly Glu Glu Gly Phe Arg Met Leu Ser Val Pro Val Pro Thr Thr
690 695 700
Tyr Ser Ala Phe Leu Asp Pro Ile Trp Thr Gln Gly Ala Asp Glu Gly
705 710 715 720
Ala Asn Val Ser Gly Gly Asp Pro Asn Val Phe Thr Trp Asp Asn Thr
725 730 735
Ser Glu Asn Gly Ser Ser Glu Asn Trp Ser Gly Leu Asn Asp Leu Thr
740 745 750
Gly Lys Ile Glu Ala Gly Thr Gly Phe Leu Val Tyr Val Tyr Lys Asn
755 760 765
Asp Glu Phe Gly Val Glu Gly Ile Phe Pro Lys Thr Leu Arg Val Ser
770 775 780
Gly Pro Ala His Gly Asp Thr Ser Pro Glu Met Asn Gln Asn Asp Gly
785 790 795 800
Gly Phe Thr Leu Leu Gly Asn Pro Phe Ser Thr Thr Ile Asp Phe Asn
805 810 815
Gln Leu Asp Lys Asn Gly Leu Thr Asp Val Ala Tyr Val Trp Asp Pro
820 825 830
Asn Ala Gly Ser Trp Arg Ser Trp Asp Ala Asp Gly Asn Ala Gly Asp
835 840 845
Leu Ala Asn Gly Leu Ile Ala Pro Phe Gln Gly Phe Phe Val Gln Ser
850 855 860
Gly Ser Gly Ser Pro Ser Val Thr Phe Asn Asp Gln Thr Lys Ser Ser
865 870 875 880
Gly Gly Val Phe Tyr Gly Lys Glu Lys Asp Glu Gln Arg Ser Ile Val
885 890 895
Arg Leu Glu Leu Ser Gly Glu Glu Leu Gln Asn Ser Ala Trp Leu Arg
900 905 910
Phe Ser Glu Ser Gly Ser Tyr Glu Ser Gln Val Arg Gly Asp Ala Leu
915 920 925
Gln Leu Glu Pro Leu Ser Gln Thr Phe Ala Lys Leu Ala Ile Glu Lys
930 935 940
Gly Gly Thr Leu Phe Asp Ile Ser His Leu Pro Val Met Asp Glu Glu
945 950 955 960
Tyr Ser Val Pro Val His Ile Ser Ala Thr Asn Ser Gly Val Tyr Thr
965 970 975
Leu Thr Ala Thr Asp Phe Glu Leu Pro Gly Asn Leu Glu Leu Thr Phe
980 985 990
His Asp His Ala Ala Gly Thr Ser Val Pro Leu Asn Ala Asp Phe Ser
995 1000 1005
Tyr Thr Ile Glu Gln Ala Gln Ala Ala Lys Ala Ala Glu Leu Ser Pro
1010 1015 1020
Leu Glu Arg Val Lys Gln Gly Pro Val Val Ala Lys Val Thr Ala Asp
1025 1030 1035 1040
Ser Pro Gln Tyr Ser Ile Thr Val Asn Ser Ser Ala Asp Ile Gly Gly
1045 1050 1055
Thr Asp Glu Leu Pro Glu Arg Val Ala Leu Arg Gln Asn Tyr Pro Asn
1060 1065 1070
Pro Phe Asn Pro Val Thr Val Ile Thr Tyr Glu Leu Pro Glu Asn Ser
1075 1080 1085
Asn Val Thr Leu Glu Val Phe Asp Met Ala Gly Arg Lys Val Ser Thr
1090 1095 1100

Leu Val Asn Glu Lys Val Ser Ala Gly Ser His Asp Val Ala Phe Asp
1105 1110 1115 1120
Ala Ser His Leu Ser Ser Gly Val Tyr Ile Tyr Arg Leu His Ala Gly
1125 1130 1135
Glu His Met Leu Thr Arg Lys Leu Thr Leu Ile Lys
1140 1145

<210> 442
<211> 1668
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 442
atgaaccgat acctgcgact ggccgcgctg accctggccc tggcgccgct ggcctacccc 60
tggggcaacc tggtgcgcgc cgccgacgcg ccgggcaaga ccgccagcgg cgtgcgctat 120
cacggcggcg acgagatcat cctccagggc tttcactgga ataccgtgcg gacctcgagc 180
aactggtacg cgacgttggc cagcatggcg ccgaccctgg ccgccgatgg cttcagtgcg 240
atctggatgc cggtgccctg gcgcgacttc tccagctgga gcgaccccgg caacggcacc 300
tcgggcggtg gcgaaggcta cttctggcac gacttcaaca agaacggccg ctatggcagc 360
gacagcctgc tcagacaggc cgccggcgcg ctcaatgcgg ccggggtcaa acccatctac 420
gacgtggtgc ccaaccacat gaaccgcggc tacccggaca aggaaatcaa cctgccggcc 480
ggccaggggc tgtggcgcca cgactgcact gacccgggca attacgccaa cgactgcgac 540
gacggcgacc gtttcatggg cggcgacgcc gacctcaata ccggccaccc gcagaactac 600
gcgatgttcc gcgacgagtt cgccaggctg cgcagcgagt atggcgccgg gggctttcgc 660
ttcgacttcg tccgcggcta cgccggcgag cgggtcgcca gctggatgag cgatgcccac 720
gacaacggct tctgcctcgg cgagctgtgg aaggcgccgg gcgagtaccc gagctgggac 780
tggcgcaacg gcgccagctg gcagcagatc ctcaaggact ggtccgatcg cgccaagtgc 840
acggtgttcg acttcgccct caaggagcgc atgcagaacg gtggcatcgc cgactggcgc 900
cacggcctca acggcaaccc cgatgcgcgc tggcgcgagg tggcggtgac cttcgtcgac 960
aaccacgaca ccggctactc gcccggcccg cacggcggcc agcaccactg gccgctgccg 1020
gatgcgcggc tgaagcaggc ctatgcctac atcctcagca gcccgggtac gccggtggtg 1080
tactggccac acatgtacga ctggggtcat ggcgacttca tccgccagtt gatccagatc 1140
cgccgtgcgg ccggagtcaa ggcggcctcg gcgatccagt tccactcggg gttctccggc 1200
ctggtggcga ccatcagcgg cagccagcag caactgctga tcgccctgga ctccaacctg 1260
tcctcgcccg gccaggtggc cagcggcgat ttcacccagg cgctcaacac cgacaacggc 1320
gcgatccgta tctggcgcag cggtcagggc ggcggggatg agccgggcaa cctggtcagc 1380
gtcaacttcc gttgcgacaa cggcgtgacc cagtggggcg acagcgtcta cgcgctgggc 1440
aacgtcgccc agctgggcaa ctggagcccg gccggcgccg tgcgcctgac cgacaccagc 1500
gcctatccaa cctggaaggg cagcattgcc ctgccggccg gtcagcaggt gcagtggaag 1560
tgcatcgtgc gcagcgaaag caacccgacc caggtcaaga cctggcagcc gggcggcaac 1620
aacagcgtga cggtggcggc aggagccagt accgcgggga gtttctag 1668

<210> 443
<211> 555
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(27)

<221> DOMAIN
<222> (45)...(395)
<223> Alpha amylase, catalytic domain

<221> DOMAIN
<222> (458)...(554)
<223> Starch binding domain

<400> 443
Met Asn Arg Tyr Leu Arg Leu Ala Ala Leu Thr Leu Ala Leu Ala Pro
1 5 10 15
Leu Ala Tyr Pro Trp Gly Asn Leu Val Arg Ala Ala Asp Ala Pro Gly

```
            20                  25                  30
Lys Thr Ala Ser Gly Val Arg Tyr His Gly Gly Asp Glu Ile Ile Leu
        35                  40                  45
Gln Gly Phe His Trp Asn Thr Val Arg Thr Ser Ser Asn Trp Tyr Ala
    50                  55                  60
Thr Leu Ala Ser Met Ala Pro Thr Leu Ala Ala Asp Gly Phe Ser Ala
65                  70                  75                  80
Ile Trp Met Pro Val Pro Trp Arg Asp Phe Ser Ser Trp Ser Asp Pro
                85                  90                  95
Gly Asn Gly Thr Ser Gly Gly Gly Glu Gly Tyr Phe Trp His Asp Phe
            100                 105                 110
Asn Lys Asn Gly Arg Tyr Gly Ser Asp Ser Leu Leu Arg Gln Ala Ala
        115                 120                 125
Gly Ala Leu Asn Ala Ala Gly Val Lys Pro Ile Tyr Asp Val Val Pro
    130                 135                 140
Asn His Met Asn Arg Gly Tyr Pro Asp Lys Glu Ile Asn Leu Pro Ala
145                 150                 155                 160
Gly Gln Gly Leu Trp Arg His Asp Cys Thr Asp Pro Gly Asn Tyr Ala
                165                 170                 175
Asn Asp Cys Asp Asp Gly Asp Arg Phe Met Gly Gly Asp Ala Asp Leu
            180                 185                 190
Asn Thr Gly His Pro Gln Asn Tyr Ala Met Phe Arg Asp Glu Phe Ala
        195                 200                 205
Arg Leu Arg Ser Glu Tyr Gly Ala Gly Gly Phe Arg Phe Asp Phe Val
    210                 215                 220
Arg Gly Tyr Ala Gly Glu Arg Val Ala Ser Trp Met Ser Asp Ala His
225                 230                 235                 240
Asp Asn Gly Phe Cys Leu Gly Glu Leu Trp Lys Ala Pro Gly Glu Tyr
                245                 250                 255
Pro Ser Trp Asp Trp Arg Asn Gly Ala Ser Trp Gln Gln Ile Leu Lys
            260                 265                 270
Asp Trp Ser Asp Arg Ala Lys Cys Thr Val Phe Asp Phe Ala Leu Lys
        275                 280                 285
Glu Arg Met Gln Asn Gly Gly Ile Ala Asp Trp Arg His Gly Leu Asn
    290                 295                 300
Gly Asn Pro Asp Ala Arg Trp Arg Glu Val Ala Val Thr Phe Val Asp
305                 310                 315                 320
Asn His Asp Thr Gly Tyr Ser Pro Gly Pro His Gly Gly Gln His His
                325                 330                 335
Trp Pro Leu Pro Asp Ala Arg Leu Lys Gln Ala Tyr Ala Tyr Ile Leu
            340                 345                 350
Ser Ser Pro Gly Thr Pro Val Val Tyr Trp Pro His Met Tyr Asp Trp
        355                 360                 365
Gly His Gly Asp Phe Ile Arg Gln Leu Ile Gln Ile Arg Arg Ala Ala
    370                 375                 380
Gly Val Lys Ala Ala Ser Ala Ile Gln Phe His Ser Gly Phe Ser Gly
385                 390                 395                 400
Leu Val Ala Thr Ile Ser Gly Ser Gln Gln Gln Leu Leu Ile Ala Leu
                405                 410                 415
Asp Ser Asn Leu Ser Ser Pro Gly Gln Val Ala Ser Gly Asp Phe Thr
            420                 425                 430
Gln Ala Leu Asn Thr Asp Asn Gly Ala Ile Arg Ile Trp Arg Ser Gly
        435                 440                 445
Gln Gly Gly Gly Asp Glu Pro Gly Asn Leu Val Ser Val Asn Phe Arg
    450                 455                 460
Cys Asp Asn Gly Val Thr Gln Trp Gly Asp Ser Val Tyr Ala Leu Gly
465                 470                 475                 480
Asn Val Ala Gln Leu Gly Asn Trp Ser Pro Ala Gly Ala Val Arg Leu
                485                 490                 495
Thr Asp Thr Ser Ala Tyr Pro Thr Trp Lys Gly Ser Ile Ala Leu Pro
            500                 505                 510
Ala Gly Gln Gln Val Gln Trp Lys Cys Ile Val Arg Ser Glu Ser Asn
        515                 520                 525
Pro Thr Gln Val Lys Thr Trp Gln Pro Gly Gly Asn Asn Ser Val Thr
    530                 535                 540
Val Ala Ala Gly Ala Ser Thr Ala Gly Ser Phe
545                 550                 555
```

<210> 444
<211> 1521
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 444
atgacaccgt ttggccagcc catgatgccg ggcgcccgga tggcagccgc caatatggcc 60
ccggtgcgtg ccgcttcggc cctgaccgat gtcagcagtg cccggcgcac ctggcatgat 120
gagatcatct attttatttt taccgatcgc tttgagaacg gtgatcgcag caatgattac 180
cgggtcaagc ccaatgaccc ctgggcctat catggcggtg accttcaggg cattatcaat 240
cgactggact atatacagaa tctgggagcc accgctctct ggatcacccc ggtgattgat 300
aaccgggatg gccccttgt ggcagatttt ggcaatggcc gcaaacagga aatctggggt 360
taccacggct actggtttaa agacttttat aaagtcgagg agcatctggg tgatatggcc 420
aagctgcgcg agctggttca gaaagcccac cagcgtggca tcaaggtgct gctggatatt 480
gtggtcaatc acacggatta tgatcatccc tttgccctgc aggcccagga gcctggtaat 540
aaataccaca gctggtttaa tcagcatggt gatatccgtg actggaatga cccctggtgg 600
gttgagaacg gccgcctggc tggcctgcct gacctgaacc agggcaatcc tgaaaccgcc 660
cgttatctga ttgatgccat gaagttctgg atcaaagaaa ccggcgtgga tggcttccgc 720
attgatacgg tcaagcacgt accccgcact ttctggcagc agtttaaccg tgaaattcgc 780
cagtttgccg gcgacgactt tttgctgctg ggtgagattt ttaccggtca ccccgaagtg 840
caggtgccct atctgcgtga aggcatgcat accgcctttg actttccgct gtactatgcc 900
atcaaagatg cctggggtca gagtcgcagc atgcgcagtc tgggggcaat ttttgccaaa 960
gaccatcttt accccgacgc caatttgctt tcgcccttta ttgataacca tgacgtgccg 1020
cgctttgtac atgaagccaa tggtcgccag cgcgatcgcc tgatggctgc tctgggcttt 1080
atctttgcga ttcgcggcat gccaagtctt tactacggta ctgaagtggc cctgcccggt 1140
ggtggggatc ccgacaaccg ccgtgatatg cagtttaacc gggaccctga gctgcaggct 1200
tatgtcaagc gccttgctga gatgcgcaag cagcagccgg ccctgcgccg tggccgccag 1260
ctggagatgt ggcaggacga tcaggtgtat gccttttcac gcctgacccc tcaggccaac 1320
gaagaggtgc tggccttttt taacaacagc tttcaacccc agcagcgccg gattgccttg 1380
cgggatgaaa gccccaataa agggcgtaac ccccgtctgg tcaatctcct gaacgcgcag 1440
gataccctgc aggtacagga tggcgcggtg aatatcaccg tgccacccat gggctatctg 1500
ttttaccgtc cggccaactg a 1521

<210> 445
<211> 506
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(24)

<221> DOMAIN
<222> (45)...(409)
<223> Alpha amylase, catalytic domain

<400> 445
Met Thr Pro Phe Gly Gln Pro Met Met Pro Gly Ala Arg Met Ala Ala
 1   5   10   15
Ala Asn Met Ala Pro Val Arg Ala Ala Ser Ala Leu Thr Asp Val Ser
            20   25   30
Ser Ala Arg Arg Thr Trp His Asp Glu Ile Ile Tyr Phe Ile Phe Thr
        35   40   45
Asp Arg Phe Glu Asn Gly Asp Arg Ser Asn Asp Tyr Arg Val Lys Pro
    50   55   60
Asn Asp Pro Trp Ala Tyr His Gly Gly Asp Leu Gln Gly Ile Ile Asn
65   70   75   80
Arg Leu Asp Tyr Ile Gln Asn Leu Gly Ala Thr Ala Leu Trp Ile Thr
                85   90   95
Pro Val Ile Asp Asn Arg Asp Gly Pro Phe Val Ala Asp Phe Gly Asn
            100   105   110
Gly Arg Lys Gln Glu Ile Trp Gly Tyr His Gly Tyr Trp Phe Lys Asp
        115   120   125

```
Phe Tyr Lys Val Glu Glu His Leu Gly Asp Met Ala Lys Leu Arg Glu
    130                 135                 140
Leu Val Gln Lys Ala His Gln Arg Gly Ile Lys Val Leu Leu Asp Ile
145                 150                 155                 160
Val Val Asn His Thr Asp Tyr Asp His Pro Phe Ala Leu Gln Ala Gln
                165                 170                 175
Glu Pro Gly Asn Lys Tyr His Ser Trp Phe Asn Gln His Gly Asp Ile
            180                 185                 190
Arg Asp Trp Asn Asp Pro Trp Trp Val Glu Asn Gly Arg Leu Ala Gly
        195                 200                 205
Leu Pro Asp Leu Asn Gln Gly Asn Pro Glu Thr Ala Arg Tyr Leu Ile
    210                 215                 220
Asp Ala Met Lys Phe Trp Ile Lys Glu Thr Gly Val Asp Gly Phe Arg
225                 230                 235                 240
Ile Asp Thr Val Lys His Val Pro Arg Thr Phe Trp Gln Gln Phe Asn
                245                 250                 255
Arg Glu Ile Arg Gln Phe Ala Gly Asp Asp Phe Leu Leu Leu Gly Glu
            260                 265                 270
Ile Phe Thr Gly His Pro Glu Val Gln Val Pro Tyr Leu Arg Glu Gly
        275                 280                 285
Met His Thr Ala Phe Asp Phe Pro Leu Tyr Tyr Ala Ile Lys Asp Ala
    290                 295                 300
Trp Gly Gln Ser Arg Ser Met Arg Ser Leu Gly Ala Ile Phe Ala Lys
305                 310                 315                 320
Asp His Leu Tyr Pro Asp Ala Asn Leu Leu Ser Pro Phe Ile Asp Asn
                325                 330                 335
His Asp Val Pro Arg Phe Val His Glu Ala Asn Gly Arg Gln Arg Asp
            340                 345                 350
Arg Leu Met Ala Ala Leu Gly Phe Ile Phe Ala Ile Arg Gly Met Pro
        355                 360                 365
Ser Leu Tyr Tyr Gly Thr Glu Val Ala Leu Pro Gly Gly Gly Asp Pro
    370                 375                 380
Asp Asn Arg Arg Asp Met Gln Phe Asn Arg Asp Pro Glu Leu Gln Ala
385                 390                 395                 400
Tyr Val Lys Arg Leu Ala Glu Met Arg Lys Gln Gln Pro Ala Leu Arg
                405                 410                 415
Arg Gly Arg Gln Leu Glu Met Trp Gln Asp Asp Gln Val Tyr Ala Phe
            420                 425                 430
Ser Arg Leu Thr Pro Gln Ala Asn Glu Glu Val Leu Ala Phe Phe Asn
        435                 440                 445
Asn Ser Phe Gln Pro Gln Gln Arg Arg Ile Ala Leu Arg Asp Glu Ser
    450                 455                 460
Pro Asn Lys Gly Arg Asn Pro Arg Leu Val Asn Leu Leu Asn Ala Gln
465                 470                 475                 480
Asp Thr Leu Gln Val Gln Asp Gly Ala Val Asn Ile Thr Val Pro Pro
                485                 490                 495
Met Gly Tyr Leu Phe Tyr Arg Pro Ala Asn
            500                 505
```

<210> 446
<211> 3537
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 446

```
gtggcctggg acacgctccc ttcgcgcgcc gacgaccagc cgggcttcca gttcacacca      60
tcgcccgcgg actggcgcac gatcagcatg taccagatca tgacggaccg gtttaacgat     120
ggcgacccgt cgaataacac cgccaatggc gtctacgacc cggtcaaccc gagtgcgatt     180
cacggcggcg actttgtggg catccggcag aagctggatt acatccagga actcggcatc     240
agcgccatct ggatttcacc ggtgcagttc tccgtgaatg cctaccacgc ctatgccccg     300
accgatttca accggatcga aacccacttc ggcagcctgc aagaactccg ggaactcatt     360
gacgactgcc acgcccgcgg catttacgtg atcatcgata tcgtcatcaa ccacatggcc     420
gacttgatcg gttccagcga ccccgggttc ccggccttca actggaatgg atacaatttg     480
aggtggtggg acccgaaccg tcgccatgcg ccgccgttcg acgacctgag ccgattccat     540
aatttcgggg aaatcgacaa ttacgacgat ccctggcagg tgctctatgg cgagttcaag     600
```

```
ccgggcggac tcgatgatat acggaccgaa gacccgcagg tgcgccagga cttgatccgc     660
attttcaagg cgctgatcga cgccaccgac gcggatggct tccgcgtgga cgccgtgaag     720
catgtggagg agggtttctt cgccgagttc atgccggcca tctacgcgca cgccgcctac     780
cgcggcaaag agaacttcct gatcttcggg gaatcgattg acgggaacga cgtgaccctc     840
tcgcggtgga cccacccgaa ttatcacttc aactccatgc tcaatttccc gatgtattac     900
aagatgatgg acgtcttcgt gcatcggcag cggacgagtc tgctcacgga gcggatcaat     960
gatctctggc gatacagcga agggtcgcgc ctgcaactcg tgaacttcct ggataaccac    1020
gacacatggc ggattatgta cgatcagaac ctggatgggg acatctggcg aatccgcccc    1080
gcactgacct ttctctatac ctcgctccag attccctgcc tctactacgg caccgaacag    1140
gccttcaacg gcggaaccga tccatggaat cgcgagaaca tgttcgacgg cggtttccag    1200
tggggcccct cggcagggga tcgcttcgac accacacatc cgctctacca gttcatccgc    1260
aagctgaacc agttgcggcg ggactatccc gctttgacgc tggggaattt cgtgcagcgc    1320
tggcaaactt cgtcggggcc gggcatctat gcctactcga agattctcgg cgaccaggaa    1380
ttgctggtgg tgctgaatac aagcgcgtcg gtgcagaact gcacgcctgc ggtttcgcgg    1440
ccgaacggta ccgtgtttgt gaacctgctc aatccatccg agcaggtcac cgtctccggt    1500
ggtacgctgg cggtcagcgt gctggggcac gaccagaaga tcttcgccga gccgcgcccg    1560
ccgctctggc tgggcaacgt gagcacctgg cccggcccca cccataacca gtgtccgggc    1620
gagttgtgca accccacgct ggaacccgat gaggagttgt gggtcgacgc cgagacccgg    1680
ccgatacggc ccggacaaac cgtgaccgtc tattatcgca ccgcggccgc cgcgacgtgg    1740
accgcccgcg aaatgctctg gcgcgaaaac acggcgcagg gatcgctctg gcatgtgaat    1800
ctgggcgggt tccccgcggg cgtgacgatc gagtattacg tgcgccgccac cgacggcgat    1860
gaagttcaat atgccaacaa cggcgggtcg aacagcttca tttccgtcgt ctatgaagca    1920
gtgccctcct ccttatccgt cgcgccatgg cccagcatcg gctgcgagcc ggtcaccttg    1980
ttgtataaac ccaacgatgg tccgctcgcc ggcgcctccg cgatctatgc gcatctcgga    2040
cgcaacggct ggaacgatat cgtcacgccc gcgcccgcgc tggaatacgt ggccgggctg    2100
gatgcgtggc gcctggtcgt cacgccgagt cccggcacgt tgcagctcga tgtcgtcttc    2160
catgatgggg ccggcgtctg ggacaacaac gatttcgaga actggttcgt ccccgtgcag    2220
gattgcgatg aagagccggg cgtcttcctg gtcacgtatc cggccgagtc gccttacctt    2280
gcagagtctt cagttgcatc aatcgatttc gagggcgtcg ccgacgggct tgccgggcaa    2340
ctggtttgga gtaacgcgct cacgggagca acgggcgttc agccgctggc acccgtctgg    2400
cgcattgacg ccatcggcct cgccgagggc cataacccga ttaccctctg ggccctgcat    2460
gacgccgggc cttccgttac gcaggcggtg gaccgcgcct tgcactacgc ggcaggcggc    2520
tggtccgacg gcggcgacct cggcggcggc tggggcggtg gctggtttct gcagtccttc    2580
ggcaccgccg gtcattttct tgccgatgca ggggtgaacg ggaatctgag gctcggagat    2640
catgcatggg ggctctggtc acaaaccgat gcccttgccg aggcggtgcg ccccttcgcg    2700
catccgctga gcgtgggcca acgtttcgag gtgcgcctcc agaacaactg gttgctcgaa    2760
gggcagcagg gggtcggcgt ggcgctgcgc gacggcgatg gtgcatcgcg cgtccagttc    2820
tacttcaacg ggggcgacga gcaatacagc gttgccgatc gcgacggcgc ccgccccagt    2880
ggcatacact ggacggatga gccacttccc ctttcgatcg aagttctgtc ggccggcgtc    2940
taccgtctta cggtcaacgg cgtcgaagtg gacggtcact tcgagggcga cctgagtgaa    3000
gtgcgcgtct ggagttataa tggaggactg acccccgact atgacttctt cttcgatgac    3060
atccgaatcg tcgctcccac tttcaccgca acggatcgcg tcgatcgcgt tgtgatccgg    3120
tcgcagagcg gaactggcga cagcaacgga gatggaatcc ccgatgcatg gtatctgcgg    3180
tatggcttta atcccttcgg accttcgatt gcgcatctgg acagcgacgg tgacggttct    3240
tccaacgggg aggagtttgt ggccgacacg aatccgaacg atccttcttc ctacttcccg    3300
aagacgatcg aggcgaatat gcagggcgcc accataacct tcacgctggg tccgccgacc    3360
tcgccggacc gcgagtacga gattctgcat tgtccggata tggtgacggg aatctggaca    3420
ccgatgaatc tacgtcgatc cggccgcgcc gaccggggtg caatgcagat cacagtgacg    3480
aaccagggcc cgcagggcgc ctaccgcacc cgcgtctcat tgccgacgga gccataa       3537

<210> 447
<211> 1178
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> DOMAIN
<222> (32)...(426)
<223> Alpha amylase, catalytic domain

<400> 447
Met Ala Trp Asp Thr Leu Pro Ser Arg Ala Asp Asp Gln Pro Gly Phe
 1               5                  10                  15
Gln Phe Thr Pro Ser Pro Ala Asp Trp Arg Thr Ile Ser Met Tyr Gln
            20                  25                  30
Ile Met Thr Asp Arg Phe Asn Asp Gly Asp Pro Ser Asn Asn Thr Ala
```

```
        35                  40                  45
Asn Gly Val Tyr Asp Pro Val Asn Pro Ser Ala Ile His Gly Gly Asp
    50                  55                  60
Phe Val Gly Ile Arg Gln Lys Leu Asp Tyr Ile Gln Glu Leu Gly Ile
65                  70                  75                  80
Ser Ala Ile Trp Ile Ser Pro Val Gln Phe Ser Val Asn Ala Tyr His
                85                  90                  95
Ala Tyr Ala Pro Thr Asp Phe Asn Arg Ile Glu Thr His Phe Gly Ser
            100                 105                 110
Leu Gln Glu Leu Arg Glu Leu Ile Asp Asp Cys His Ala Arg Gly Ile
        115                 120                 125
Tyr Val Ile Ile Asp Ile Val Ile Asn His Met Ala Asp Leu Ile Gly
    130                 135                 140
Ser Ser Asp Pro Gly Phe Pro Ala Phe Asn Trp Asn Gly Tyr Asn Leu
145                 150                 155                 160
Arg Trp Trp Asp Pro Asn Arg Arg His Ala Pro Pro Phe Asp Asp Leu
                165                 170                 175
Ser Arg Phe His Asn Phe Gly Glu Ile Asp Asn Tyr Asp Asp Pro Trp
            180                 185                 190
Gln Val Leu Tyr Gly Glu Phe Lys Pro Gly Gly Leu Asp Asp Ile Arg
        195                 200                 205
Thr Glu Asp Pro Gln Val Arg Gln Asp Leu Ile Arg Ile Phe Lys Ala
    210                 215                 220
Leu Ile Asp Ala Thr Asp Ala Asp Gly Phe Arg Val Asp Ala Val Lys
225                 230                 235                 240
His Val Glu Glu Gly Phe Phe Ala Glu Phe Met Pro Ala Ile Tyr Ala
                245                 250                 255
His Ala Ala Tyr Arg Gly Lys Glu Asn Phe Leu Ile Phe Gly Glu Ser
            260                 265                 270
Ile Asp Gly Asn Asp Val Thr Leu Ser Arg Trp Thr His Pro Asn Tyr
        275                 280                 285
His Phe Asn Ser Met Leu Asn Phe Pro Met Tyr Tyr Lys Met Met Asp
    290                 295                 300
Val Phe Val His Arg Gln Arg Thr Ser Leu Leu Thr Glu Arg Ile Asn
305                 310                 315                 320
Asp Leu Trp Arg Tyr Ser Glu Gly Ser Arg Leu Gln Leu Val Asn Phe
                325                 330                 335
Leu Asp Asn His Asp Thr Trp Arg Ile Met Tyr Asp Gln Asn Leu Asp
            340                 345                 350
Gly Asp Ile Trp Arg Ile Arg Pro Ala Leu Thr Phe Leu Tyr Thr Ser
        355                 360                 365
Leu Gln Ile Pro Cys Leu Tyr Tyr Gly Thr Glu Gln Ala Phe Asn Gly
    370                 375                 380
Gly Thr Asp Pro Trp Asn Arg Glu Asn Met Phe Asp Gly Gly Phe Gln
385                 390                 395                 400
Trp Gly Pro Ser Ala Gly Asp Arg Phe Asp Thr Thr His Pro Leu Tyr
                405                 410                 415
Gln Phe Ile Arg Lys Leu Asn Gln Leu Arg Arg Asp Tyr Pro Ala Leu
            420                 425                 430
Thr Leu Gly Asn Phe Val Gln Arg Trp Gln Thr Ser Ser Gly Pro Gly
        435                 440                 445
Ile Tyr Ala Tyr Ser Lys Ile Leu Gly Asp Gln Glu Leu Leu Val Val
    450                 455                 460
Leu Asn Thr Ser Ala Ser Val Gln Asn Cys Thr Pro Ala Val Ser Arg
465                 470                 475                 480
Pro Asn Gly Thr Val Phe Val Asn Leu Leu Asn Pro Ser Glu Gln Val
                485                 490                 495
Thr Val Ser Gly Gly Thr Leu Ala Val Ser Val Leu Gly His Asp Gln
            500                 505                 510
Lys Ile Phe Ala Glu Pro Arg Pro Pro Leu Trp Leu Gly Asn Val Ser
        515                 520                 525
Thr Trp Pro Gly Pro Thr His Asn Gln Cys Pro Gly Glu Leu Cys Asn
    530                 535                 540
Pro Thr Leu Glu Pro Asp Glu Glu Leu Trp Val Asp Ala Glu Thr Arg
545                 550                 555                 560
Pro Ile Arg Pro Gly Gln Thr Val Thr Val Tyr Tyr Arg Thr Ala Ala
                565                 570                 575
Ala Ala Thr Trp Thr Ala Arg Glu Met Leu Trp Arg Glu Asn Thr Ala
```

580 585 590
Gln Gly Ser Leu Trp His Val Asn Leu Gly Gly Phe Pro Ala Gly Val
595 600 605
Thr Ile Glu Tyr Tyr Val Arg Ala Thr Asp Gly Asp Glu Val Gln Tyr
610 615 620
Ala Asn Asn Gly Gly Ser Asn Ser Phe Ile Ser Val Val Tyr Glu Ala
625 630 635 640
Val Pro Ser Ser Leu Ser Val Ala Pro Trp Pro Ser Ile Gly Cys Glu
645 650 655
Pro Val Thr Leu Leu Tyr Lys Pro Asn Asp Gly Pro Leu Ala Gly Ala
660 665 670
Ser Ala Ile Tyr Ala His Leu Gly Arg Asn Gly Trp Asn Asp Ile Val
675 680 685
Thr Pro Ala Pro Ala Leu Glu Tyr Val Ala Gly Leu Asp Ala Trp Arg
690 695 700
Leu Val Val Thr Pro Ser Pro Gly Thr Leu Gln Leu Asp Val Val Phe
705 710 715 720
His Asp Gly Ala Gly Val Trp Asp Asn Asn Asp Phe Glu Asn Trp Phe
725 730 735
Val Pro Val Gln Asp Cys Asp Glu Glu Pro Gly Val Phe Leu Val Thr
740 745 750
Tyr Pro Ala Glu Ser Pro Tyr Leu Ala Glu Ser Ser Val Ala Ser Ile
755 760 765
Asp Phe Glu Gly Val Ala Asp Gly Leu Ala Gly Gln Leu Val Trp Ser
770 775 780
Asn Ala Leu Thr Gly Ala Thr Gly Val Gln Pro Leu Ala Pro Val Trp
785 790 795 800
Arg Ile Asp Ala Ile Gly Leu Ala Glu Gly His Asn Pro Ile Thr Leu
805 810 815
Trp Ala Leu His Asp Ala Gly Pro Ser Val Thr Gln Ala Val Asp Arg
820 825 830
Ala Leu His Tyr Ala Ala Gly Gly Trp Ser Asp Gly Gly Asp Leu Gly
835 840 845
Gly Gly Trp Gly Gly Gly Trp Phe Leu Gln Ser Phe Gly Thr Ala Gly
850 855 860
His Phe Leu Ala Asp Ala Gly Val Asn Gly Asn Leu Arg Leu Gly Asp
865 870 875 880
His Ala Trp Gly Leu Trp Ser Gln Thr Asp Ala Leu Ala Glu Ala Val
885 890 895
Arg Pro Phe Ala His Pro Leu Ser Val Gly Gln Arg Phe Glu Val Arg
900 905 910
Leu Gln Asn Asn Trp Leu Leu Glu Gly Gln Gln Gly Val Gly Val Ala
915 920 925
Leu Arg Asp Gly Asp Gly Ala Ser Arg Val Gln Phe Tyr Phe Asn Gly
930 935 940
Gly Asp Glu Gln Tyr Ser Val Ala Asp Arg Asp Gly Ala Arg Pro Ser
945 950 955 960
Gly Ile His Trp Thr Asp Glu Pro Leu Pro Leu Ser Ile Glu Val Leu
965 970 975
Ser Ala Gly Val Tyr Arg Leu Thr Val Asn Gly Val Glu Val Asp Gly
980 985 990
His Phe Glu Gly Asp Leu Ser Glu Val Arg Val Trp Ser Tyr Asn Gly
995 1000 1005
Gly Leu Thr Pro Asp Tyr Asp Phe Phe Phe Asp Asp Ile Arg Ile Val
1010 1015 1020
Ala Pro Thr Phe Thr Ala Thr Asp Arg Val Asp Arg Val Val Ile Arg
1025 1030 1035 1040
Ser Gln Ser Gly Thr Gly Asp Ser Asn Gly Asp Gly Ile Pro Asp Ala
1045 1050 1055
Trp Tyr Leu Arg Tyr Gly Phe Asn Pro Phe Gly Pro Ser Ile Ala His
1060 1065 1070
Leu Asp Ser Asp Gly Asp Gly Ser Ser Asn Gly Glu Glu Phe Val Ala
1075 1080 1085
Asp Thr Asn Pro Asn Asp Pro Ser Ser Tyr Phe Pro Lys Thr Ile Glu
1090 1095 1100
Ala Asn Met Gln Gly Ala Thr Ile Thr Phe Thr Leu Gly Pro Pro Thr
1105 1110 1115 1120
Ser Pro Asp Arg Glu Tyr Glu Ile Leu His Cys Pro Asp Met Val Thr

```
                1125                1130                1135
Gly Ile Trp Thr Pro Met Asn Leu Arg Arg Ser Gly Arg Ala Asp Arg
            1140                1145                1150
Gly Ala Met Gln Ile Thr Val Thr Asn Gln Gly Pro Gln Gly Ala Tyr
        1155                1160                1165
Arg Thr Arg Val Ser Leu Pro Thr Glu Pro
    1170                1175
```

<210> 448
<211> 1575
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 448

```
atgagattaa ttatgaaaaa aatgatcatc ctaattacgc tcgcttgggt ttttaccggc      60
tgcgagagct caaagtttaa agaagagagc aactatacct caagctactc tgccgaatgg     120
ccaatggcga cgagctatga gatttttgtt cgctcatttg ctgatagcaa tggtgacggt     180
attggagata taaaaggaat gacagaaaag ttaccttatt tgcaagacct gggaattaaa     240
ggagtatggt taatgccgat tcatccctca ccttcttacc ataaatatga tgtgacggac     300
tataaaggcg tgcatgaaga ctatggtacc ctagaggatt tcaaaaagtt tgtaaagcga     360
gcgcatgagc atgatattaa agtagtgata gattttgtga tcaatcatac agcgagggag     420
cacccttggt ttcaagagtc gatgaagggg cctgataatc cttatcgcga ctactatgtg     480
tgggcgaaaa aggaggatat agaagatgaa atcaataaaa aggaaacggc attagactct     540
gacaatatcg agcagtggca tgaggcaccc ggcaatgagg agctatacta cggctacttt     600
tggggtggta tgccagattt gaattttgac aatcctgaag tgaagcgtga aatatttgat     660
gcgggaaggt tttggttgga agaagttggt gtggatggct ttcgattgga tgctgcaaga     720
catatttttc cagacgatcg ggcagctgac aatcatgctt tttgggtaga atttcgcaat     780
gagatgaaaa gagtaaaccc tgatgtttat ctagttggtg aagtatgggc tgatatggaa     840
acagtagcac cttatttgag tggtttgccc gcacttttca attttgatat gagctataaa     900
atcgtggagg cggtgcagcg tcaggccaat gaagaaatag cggagtatca tgcaaaggtg     960
atgaactatt atcaagggat caaccctaat tatattgatg ccaccttcat taccaatcac    1020
gatcagccac ggagcatgag tcagtttgag ggggatgtga ataaagggaa gctagctgct    1080
tcactactat ttacactgcc tggcgcacct tatttgtact atggtgaaga aattgggatg    1140
ttggggccaa agcctgatcc tgaaatcaga gagccgtttt tatggaaaga aggtgagaag    1200
gatgtttacc gcacgagttg gatagaggcg aaaaactcta atgacgaaac agtaattcca    1260
gccaaagtac agaaagaaga tccggaaagc ttattgaacc actacaagag atgggtggct    1320
ctcagacata aaagtgaggc gcttagcgaa ggccgaatcg tgcctgtagc cacagataat    1380
caacaaatgg ctgtttggga gcgagtaagt aaaaatgaaa gggtactcgt gattcataat    1440
ctttcagata gtcctcaaac tatatcagag gaggtaagcc gtaaatatga tttggctaat    1500
aaaaagatga tatggggaag tgtacaaaac aatgagctaa gtcctagagg gtctttagtc    1560
tttgtgatta actaa                                                     1575
```

<210> 449
<211> 524
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> DOMAIN
<222> (47)...(442)
<223> Alpha amylase, catalytic domain

<221> SIGNAL
<222> (1)...(23)

<400> 449

```
Met Arg Leu Ile Met Lys Lys Met Ile Ile Leu Ile Thr Leu Ala Trp
 1               5                  10                  15
Val Phe Thr Gly Cys Glu Ser Ser Lys Phe Lys Glu Glu Ser Asn Tyr
            20                  25                  30
Thr Ser Ser Tyr Ser Ala Glu Trp Pro Met Ala Thr Ser Tyr Glu Ile
        35                  40                  45
Phe Val Arg Ser Phe Ala Asp Ser Asn Gly Asp Gly Ile Gly Asp Ile
```

```
    50                  55                  60
Lys Gly Met Thr Glu Lys Leu Pro Tyr Leu Gln Asp Leu Gly Ile Lys
65                  70                  75                  80
Gly Val Trp Leu Met Pro Ile His Pro Ser Pro Ser Tyr His Lys Tyr
                85                  90                  95
Asp Val Thr Asp Tyr Lys Gly Val His Glu Asp Tyr Gly Thr Leu Glu
            100                 105                 110
Asp Phe Lys Lys Phe Val Lys Arg Ala His Glu His Asp Ile Lys Val
        115                 120                 125
Val Ile Asp Phe Val Ile Asn His Thr Ala Arg Glu His Pro Trp Phe
    130                 135                 140
Gln Glu Ser Met Lys Gly Pro Asp Asn Pro Tyr Arg Asp Tyr Tyr Val
145                 150                 155                 160
Trp Ala Lys Lys Glu Asp Ile Glu Asp Glu Ile Asn Lys Lys Glu Thr
                165                 170                 175
Ala Leu Asp Ser Asp Asn Ile Glu Gln Trp His Glu Ala Pro Gly Asn
            180                 185                 190
Glu Glu Leu Tyr Tyr Gly Tyr Phe Trp Gly Gly Met Pro Asp Leu Asn
        195                 200                 205
Phe Asp Asn Pro Glu Val Lys Arg Glu Ile Phe Asp Ala Gly Arg Phe
    210                 215                 220
Trp Leu Glu Glu Val Gly Val Asp Gly Phe Arg Leu Asp Ala Ala Arg
225                 230                 235                 240
His Ile Phe Pro Asp Asp Arg Ala Ala Asp Asn His Ala Phe Trp Val
                245                 250                 255
Glu Phe Arg Asn Glu Met Lys Arg Val Asn Pro Asp Val Tyr Leu Val
            260                 265                 270
Gly Glu Val Trp Ala Asp Met Glu Thr Val Ala Pro Tyr Leu Ser Gly
        275                 280                 285
Leu Pro Ala Leu Phe Asn Phe Asp Met Ser Tyr Lys Ile Val Glu Ala
    290                 295                 300
Val Gln Arg Gln Ala Asn Glu Glu Ile Ala Glu Tyr His Ala Lys Val
305                 310                 315                 320
Met Asn Tyr Tyr Gln Gly Ile Asn Pro Asn Tyr Ile Asp Ala Thr Phe
                325                 330                 335
Ile Thr Asn His Asp Gln Pro Arg Ser Met Ser Gln Phe Glu Gly Asp
            340                 345                 350
Val Asn Lys Gly Lys Leu Ala Ala Ser Leu Leu Phe Thr Leu Pro Gly
        355                 360                 365
Ala Pro Tyr Leu Tyr Tyr Gly Glu Glu Ile Gly Met Leu Gly Pro Lys
    370                 375                 380
Pro Asp Pro Glu Ile Arg Glu Pro Phe Leu Trp Lys Glu Gly Glu Lys
385                 390                 395                 400
Asp Val Tyr Arg Thr Ser Trp Ile Glu Ala Lys Asn Ser Asn Asp Glu
                405                 410                 415
Thr Val Ile Pro Ala Lys Val Gln Lys Glu Asp Pro Glu Ser Leu Leu
            420                 425                 430
Asn His Tyr Lys Arg Trp Val Ala Leu Arg His Lys Ser Glu Ala Leu
        435                 440                 445
Ser Glu Gly Arg Ile Val Pro Val Ala Thr Asp Asn Gln Gln Met Ala
    450                 455                 460
Val Trp Glu Arg Val Ser Lys Asn Glu Arg Val Leu Val Ile His Asn
465                 470                 475                 480
Leu Ser Asp Ser Pro Gln Thr Ile Ser Glu Glu Val Ser Arg Lys Tyr
                485                 490                 495
Asp Leu Ala Asn Lys Lys Met Ile Trp Gly Ser Val Gln Asn Asn Glu
            500                 505                 510
Leu Ser Pro Arg Gly Ser Leu Val Phe Val Ile Asn
        515                 520

<210> 450
<211> 1890
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample
```

<400> 450
atgaatgatt caatcaatct ttataacttc ttcccgtaca atagaccaat gtcaatcaat 60
aaaacaaaca ctatgaaaca aatgatcaac tggctaggta gtctcgcgct actcatgttg 120
ctattgagtt gtggtgaggc tactgagtcc tcatccctcg ccaatcaaaa aggagaaaaa 180
gagcgctcag atgaaggctc tgatgataag aaagtaattt atcaagtgtt tacacgaaca 240
tttggtaatc agaattctac taacaagatg cacggtactg ttgaggaaaa tggcacgggt 300
aagttcaatg acttcactct agacgcactc gaaggcatca aggagcttgg cgtaactcac 360
atttggtata caggagtgat tgctcatgct accgccacag atttcagtga gtatggggta 420
gcagtggata atccgagagt gatcaagggg cgagcgggta gtccttatgc tatcaaggat 480
tactatgacg taaacccgga tctagcggaa gatgtagata agcgaatgga tgaattcgaa 540
gatatggtca agcgtacgca tgagcgtggg atgaaggtga tcatagactt tgtgcctaat 600
cacgtagcgc gttcctacca aggtcagcag cagcctgagg gtatcaccgc tttgggcgct 660
gaggatgata ctagcaaagc atttgatgta aataatcact tttactactt accgggtcag 720
gattttcaag tgccggcagg ctatcgtccg ctggggcttt ctgccaagga gatggaagac 780
ttacctgctt atgaggagtc acctgctaag gcaacgggca atgatgtatt tagtgcaagt 840
ccatcagtga atgactggtt tgagacgatc aagttgaact atggtgtgga ttaccaaaat 900
ggtggtaagc agcatttttc agctactcag atccctgata cttgggagcg catgaaggag 960
atcttactct tttgggcagg taagggcgtg gatggctttc gttgtgatat ggctgagatg 1020
gtgccggtgg agttttggga atgggtgagt gcggcattga aagaagagta tcctgatatc 1080
ctattaattg cagagattta caatcccgat gcttataagc cttacatcaa acaaggaggc 1140
ttcgattatt tatatgacaa agtacaaacc tacgacagcc tgcgcaatgt gatgcaagat 1200
agaggtacta cgcggggtat ttcggcatca ttggaagcct tggtagatgt ggagccacat 1260
atgttgcgat ttctcgaaaa ccatgatgag cagcgaatag cctcttcatt tttcgcgggg 1320
gatgcttttg ccggtcttcc tgcactggcc attagcaact tggctagtgg tagtcccttc 1380
atgctttatt ttggtcagga agttggggaa cccgccactg aagctgctgg ctttagtgag 1440
gcggacggta gaagtacaat ttttgactac ttcggtatac ctgagcatca gaagtggatg 1500
aatgaaggtg cctttgatgg tggacagctt tcagaagatc agcagcggct tagagcattg 1560
taccaagaga ttatccgagt atctacactc agtcgattgg caaactcaga gccttttact 1620
tgcagattag accaagtgga gtcctcatcc cccgaaaaaa tacattgctt ttctagagaa 1680
aatgaggaag aagagtggta tgtggtcacc aacttttctg atcaatcgac tgaggtgagt 1740
gtgaaagacc taagtgcaga caaggctcag atgatgtttt actggagaaa agagcttagc 1800
gagctgcctg aggaaattga aatagaggat ggaaagttaa agtgggaatt agacccttat 1860
gccatcttgg tgttttcttt ggctaagtaa 1890

<210> 451
<211> 629
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> DOMAIN
<222> (75)...(496)
<223> Alpha amylase, catalytic domain

<221> SIGNAL
<222> (1)...(49)

<400> 451
Met Asn Asp Ser Ile Asn Leu Tyr Asn Phe Phe Pro Tyr Asn Arg Pro
 1   5   10   15
Met Ser Ile Asn Lys Thr Asn Thr Met Lys Gln Met Ile Asn Trp Leu
            20   25   30
Gly Ser Leu Ala Leu Leu Met Leu Leu Leu Ser Cys Gly Glu Ala Thr
        35   40   45
Glu Ser Ser Ser Leu Ala Asn Gln Lys Gly Glu Lys Glu Arg Ser Asp
    50   55   60
Glu Gly Ser Asp Asp Lys Lys Val Ile Tyr Gln Val Phe Thr Arg Thr
65   70   75   80
Phe Gly Asn Gln Asn Ser Thr Asn Lys Met His Gly Thr Val Glu Glu
                85   90   95
Asn Gly Thr Gly Lys Phe Asn Asp Phe Thr Leu Asp Ala Leu Glu Gly
            100   105   110
Ile Lys Glu Leu Gly Val Thr His Ile Trp Tyr Thr Gly Val Ile Ala
        115   120   125
His Ala Thr Ala Thr Asp Phe Ser Glu Tyr Gly Val Ala Val Asp Asn
    130   135   140

```
Pro Arg Val Ile Lys Gly Arg Ala Gly Ser Pro Tyr Ala Ile Lys Asp
145                 150                 155                 160
Tyr Tyr Asp Val Asn Pro Asp Leu Ala Glu Asp Val Asp Lys Arg Met
                165                 170                 175
Asp Glu Phe Glu Asp Met Val Lys Arg Thr His Glu Arg Gly Met Lys
            180                 185                 190
Val Ile Ile Asp Phe Val Pro Asn His Val Ala Arg Ser Tyr Gln Gly
        195                 200                 205
Gln Gln Gln Pro Glu Gly Ile Thr Ala Leu Gly Ala Glu Asp Asp Thr
    210                 215                 220
Ser Lys Ala Phe Asp Val Asn Asn His Phe Tyr Tyr Leu Pro Gly Gln
225                 230                 235                 240
Asp Phe Gln Val Pro Ala Gly Tyr Arg Pro Leu Gly Leu Ser Ala Lys
                245                 250                 255
Glu Met Glu Asp Leu Pro Ala Tyr Glu Glu Ser Pro Ala Lys Ala Thr
            260                 265                 270
Gly Asn Asp Val Phe Ser Ala Ser Pro Ser Val Asn Asp Trp Phe Glu
        275                 280                 285
Thr Ile Lys Leu Asn Tyr Gly Val Asp Tyr Gln Asn Gly Gly Lys Gln
    290                 295                 300
His Phe Ser Ala Thr Gln Ile Pro Asp Thr Trp Glu Arg Met Lys Glu
305                 310                 315                 320
Ile Leu Leu Phe Trp Ala Gly Lys Gly Val Asp Gly Phe Arg Cys Asp
                325                 330                 335
Met Ala Glu Met Val Pro Val Glu Phe Trp Glu Trp Val Ser Ala Ala
            340                 345                 350
Leu Lys Glu Glu Tyr Pro Asp Ile Leu Leu Ile Ala Glu Ile Tyr Asn
        355                 360                 365
Pro Asp Ala Tyr Lys Pro Tyr Ile Lys Gln Gly Gly Phe Asp Tyr Leu
    370                 375                 380
Tyr Asp Lys Val Gln Thr Tyr Asp Ser Leu Arg Asn Val Met Gln Asp
385                 390                 395                 400
Arg Gly Thr Thr Arg Gly Ile Ser Ala Ser Leu Glu Ala Leu Val Asp
                405                 410                 415
Val Glu Pro His Met Leu Arg Phe Leu Glu Asn His Asp Glu Gln Arg
            420                 425                 430
Ile Ala Ser Ser Phe Phe Ala Gly Asp Ala Phe Ala Gly Leu Pro Ala
        435                 440                 445
Leu Ala Ile Ser Asn Leu Ala Ser Gly Ser Pro Phe Met Leu Tyr Phe
    450                 455                 460
Gly Gln Glu Val Gly Glu Pro Ala Thr Glu Ala Ala Gly Phe Ser Glu
465                 470                 475                 480
Ala Asp Gly Arg Ser Thr Ile Phe Asp Tyr Phe Gly Ile Pro Glu His
                485                 490                 495
Gln Lys Trp Met Asn Glu Gly Ala Phe Asp Gly Gly Gln Leu Ser Glu
            500                 505                 510
Asp Gln Gln Arg Leu Arg Ala Leu Tyr Gln Glu Ile Ile Arg Val Ser
        515                 520                 525
Thr Leu Ser Arg Leu Ala Asn Ser Glu Pro Phe Thr Cys Arg Leu Asp
    530                 535                 540
Gln Val Glu Ser Ser Ser Pro Glu Lys Ile His Cys Phe Ser Arg Glu
545                 550                 555                 560
Asn Glu Glu Glu Glu Trp Tyr Val Val Thr Asn Phe Ser Asp Gln Ser
                565                 570                 575
Thr Glu Val Ser Val Lys Asp Leu Ser Ala Asp Lys Ala Gln Met Met
            580                 585                 590
Phe Tyr Trp Arg Lys Glu Leu Ser Glu Leu Pro Glu Glu Ile Glu Ile
        595                 600                 605
Glu Asp Gly Lys Leu Lys Trp Glu Leu Asp Pro Tyr Ala Ile Leu Val
    610                 615                 620
Phe Ser Leu Ala Lys
625
```

<210> 452
<211> 1734
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 452
gtgatgcaac tgaatccctg gttttctacc acgcttaagg ccgctggcct ggcgaccgct 60
ctggctgcgg tcagcgcttg ccagccggca tccgagtcgg cttatcaagc cccaactgca 120
gcgcctgtcg agcagaccga agcgccctgg tgggacgatg cggttttcta ccaaatctgg 180
ccgcgtagtt tttacgatac cagcggtgat ggccatggcg attttaatgg catgaccgcc 240
aaactgcctt atttgtctga gctaggcatt gatgccatct ggctaacgcc gatgtttgag 300
gcgccgtctt atcatggcta tgactttacc gagttttatc aggtggaatc ggactatggc 360
accatggccg acttcgagcg ctttttgcaa caagctgaag agcgcggtat ccgggtgatt 420
ctggatctgg tgatcaacca catttcccgt gagcaccagt ggtttaagca gtcagcccag 480
cgcattgagc cgtacagcga ttactttatc tggcgggaaa gcctgccaga agccggcaag 540
ggctggggcc atgcctggtc ggataacgat atgcccgaag cggtctggca ttttgatgag 600
acccggggcg agtactacta cgccgccttt ggtgccaccc agccggatct gaatctggag 660
cacccggatg tgatcgaaga aatgaaacgg atggctaaat tctggctgga taaaggggtc 720
gatggcttcc ggctcgatgc ggtgcgcttt gtgatagagc gcggcccgaa aggccaggcc 780
gacttgcccg agaccattga ttactggaaa gacttcaacg cctttgtcaa aagcgttaac 840
cccgaagcct acttagtggg cgaagcctgg gttgatattc cggtggcagc caaatactgg 900
ggtgaaggca aagggctgga tcagggcttt gattttgagt ttggttacat cgtgctggat 960
atgctgcagc aagggctggc cacggaagcc cagtttggca ccatgagcca gcagtccggt 1020
caggctgaag atggcgcttt gcatccactg gcgcagaatg tgcagcgccg cctggagtcg 1080
ggcgcaccgc tggccttctt cacgcccttt ttaaccaatc acgatcagga cagagttggc 1140
tttcaattgc tggaaagtca ggccaaagcc aaacaagcgg cggcgctgtt atttagctcg 1200
ccgggcacca agtacatcta ctacggtgaa gaaattggca tgacccagct aaaaagtggc 1260
catgatgttt acaagcgggc gccaatgttg tgggatcaag ggcatcaggc cggctttacc 1320
caggcagaaa aggcttgggt agagcaaaaa gatttattca gccgccattt tgatgcctgg 1380
tggccggatt tcctgacgca gcagttggca gtcggcgatc ggacagtgga agcacaacgt 1440
gcacagccgg actccttgtg gaacttgcac cgtctgctga tcgagctgca aaagcagcgg 1500
ccggaaatgc gtttggctgg ctcgtaccag gtgcaatcgc cacagcaggg tattctgtcg 1560
gtgacgcgtg agctgggcga gcaaagcacg gtgtttgtgc tgaatctggc tggtgaaatg 1620
gctgatattg ctgctgtgct ggcagcgcac agcaatctgt cagtaagttg gggctttgat 1680
ttggatggta atttgttggc tatagatggt ttgttgattt tgcaaagtca gtag 1734

<210> 453
<211> 577
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> DOMAIN
<222> (58)...(440)
<223> Alpha amylase, catalytic domain

<221> SIGNAL
<222> (1)...(34)

<400> 453
Met Met Gln Leu Asn Pro Trp Phe Ser Thr Thr Leu Lys Ala Ala Gly
 1   5   10   15
Leu Ala Thr Ala Leu Ala Ala Val Ser Ala Cys Gln Pro Ala Ser Glu
            20   25   30
Ser Ala Tyr Gln Ala Pro Thr Ala Ala Pro Val Glu Gln Thr Glu Ala
        35   40   45
Pro Trp Trp Asp Asp Ala Val Phe Tyr Gln Ile Trp Pro Arg Ser Phe
    50   55   60
Tyr Asp Thr Ser Gly Asp Gly His Gly Asp Phe Asn Gly Met Thr Ala
65   70   75   80
Lys Leu Pro Tyr Leu Ser Glu Leu Gly Ile Asp Ala Ile Trp Leu Thr
                85   90   95
Pro Met Phe Glu Ala Pro Ser Tyr His Gly Tyr Asp Phe Thr Glu Phe
            100   105   110
Tyr Gln Val Glu Ser Asp Tyr Gly Thr Met Ala Asp Phe Glu Arg Phe
        115   120   125
Leu Gln Gln Ala Glu Glu Arg Gly Ile Arg Val Ile Leu Asp Leu Val
    130   135   140

Ile Asn His Ile Ser Arg Glu His Gln Trp Phe Lys Gln Ser Ala Gln
145 150 155 160
Arg Ile Glu Pro Tyr Ser Asp Tyr Phe Ile Trp Arg Glu Ser Leu Pro
165 170 175
Glu Ala Gly Lys Gly Trp Gly His Ala Trp Ser Asp Asn Asp Met Pro
180 185 190
Glu Ala Val Trp His Phe Asp Glu Thr Arg Gly Glu Tyr Tyr Tyr Ala
195 200 205
Ala Phe Gly Ala Thr Gln Pro Asp Leu Asn Leu Glu His Pro Asp Val
210 215 220
Ile Glu Glu Met Lys Arg Met Ala Lys Phe Trp Leu Asp Lys Gly Val
225 230 235 240
Asp Gly Phe Arg Leu Asp Ala Val Arg Phe Val Ile Glu Arg Gly Pro
245 250 255
Lys Gly Gln Ala Asp Leu Pro Glu Thr Ile Asp Tyr Trp Lys Asp Phe
260 265 270
Asn Ala Phe Val Lys Ser Val Asn Pro Glu Ala Tyr Leu Val Gly Glu
275 280 285
Ala Trp Val Asp Ile Pro Val Ala Ala Lys Tyr Trp Gly Glu Gly Lys
290 295 300
Gly Leu Asp Gln Gly Phe Asp Phe Glu Phe Gly Tyr Ile Val Leu Asp
305 310 315 320
Met Leu Gln Gln Gly Leu Ala Thr Glu Ala Gln Phe Gly Thr Met Ser
325 330 335
Gln Gln Ser Gly Gln Ala Glu Asp Gly Ala Leu His Pro Leu Ala Gln
340 345 350
Asn Val Gln Arg Arg Leu Glu Ser Gly Ala Pro Leu Ala Phe Phe Thr
355 360 365
Pro Phe Leu Thr Asn His Asp Gln Asp Arg Val Gly Phe Gln Leu Leu
370 375 380
Glu Ser Gln Ala Lys Ala Lys Gln Ala Ala Ala Leu Leu Phe Ser Ser
385 390 395 400
Pro Gly Thr Lys Tyr Ile Tyr Tyr Gly Glu Glu Ile Gly Met Thr Gln
405 410 415
Leu Lys Ser Gly His Asp Val Tyr Lys Arg Ala Pro Met Leu Trp Asp
420 425 430
Gln Gly His Gln Ala Gly Phe Thr Gln Ala Glu Lys Ala Trp Val Glu
435 440 445
Gln Lys Asp Leu Phe Ser Arg His Phe Asp Ala Trp Trp Pro Asp Phe
450 455 460
Leu Thr Gln Gln Leu Ala Val Gly Asp Arg Thr Val Glu Ala Gln Arg
465 470 475 480
Ala Gln Pro Asp Ser Leu Trp Asn Leu His Arg Leu Leu Ile Glu Leu
485 490 495
Gln Lys Gln Arg Pro Glu Met Arg Leu Ala Gly Ser Tyr Gln Val Gln
500 505 510
Ser Pro Gln Gln Gly Ile Leu Ser Val Thr Arg Glu Leu Gly Glu Gln
515 520 525
Ser Thr Val Phe Val Leu Asn Leu Ala Gly Glu Met Ala Asp Ile Ala
530 535 540
Ala Val Leu Ala Ala His Ser Asn Leu Ser Val Ser Trp Gly Phe Asp
545 550 555 560
Leu Asp Gly Asn Leu Leu Ala Ile Asp Gly Leu Leu Ile Leu Gln Ser
565 570 575
Gln

<210> 454
<211> 1341
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 454
ttggatctgt tagaatataa gaatacaatt caacggagac aaacgatgac tgaccgcaag 60
ttgttattta tagttgcgac cgtgatactg gcagttctgg tgtcgttttc ggcagagctc 120

gaagcaagag tcaagcctac caatgctgtt atgctccagg gattcggctg gaactctcag 180
tcccgaggca ccccaagcaa atggtacaat ttgattggcg acagggccaa agacatagcc 240
gatctaggtg taacgatggt ctggtttccg ccggtgagtc gctctgtgag tcctcagggc 300
tatcttccgg gcgactggta tgatctcgga acagagaagg acccgacgtt ttacggccat 360
aaggatagcc ttttgtacgc tttaagcgag ctgaaaaagc acggtatttc tccgatcgcc 420
gacatagtca taaaccacag gtgcgccagc catcaggaca agaacgggat ttggaacatc 480
taccattttc cgtcaggaaa ggcgcgctgg gagcaatggg cgatctgccg cggccagttt 540
ggtggacagg gaaaccctga cacaggcgag agctaccatg cggctccaga cgttgatcat 600
acgaacgtca gggttcagcg ggatataatc gagtggatga actggctgaa atctcttggc 660
tttgaaggtt ggcgttatga ctactcaaaa ggctaccatc caaggtatac tgccctttat 720
gaccggtcta ccaggcccgt attgtctgta ggtgagatat ggacaaacat gtctttcaac 780
ggctcttatc ttaacccaaa tcaggacgcc caccgacagc agctctgcga ctggcttgac 840
ggggccggca actatgccac cgcctttgac ttcactacga aggggattct tcaggtagcc 900
gttaaaggtg agtactggcg actacgcgac aaggacggca aggcgtcggg attaataggc 960
tggtggcccg aacgcgcagt cacatttgtc gataatcacg ataccggttc ccagcaaagc 1020
cactggccat ttcccggcaa caaggttatg caggggtacg catatattct tacacatcca 1080
ggcataccct gcattttctg ggagcatgtc tacgactgga aactacgtga accgatcaaa 1140
aggcttgtcc agatcaggaa gaaatatgga ttgcacagca agagcagtct taagatcgtt 1200
agagccgagc agaatcttta tgctgcgatc attgacggaa aggtagcctt gaagcttggc 1260
agcgctgact ggtcaccgcc aggcaatgat tttgaacttc tcacacatgg cgaccagttt 1320
gcagtttggg gcagaaaata a 1341

<210> 455
<211> 446
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(37)

<221> DOMAIN
<222> (49)...(386)
<223> Alpha amylase, catalytic domain

<400> 455
Met Asp Leu Leu Glu Tyr Lys Asn Thr Ile Gln Arg Arg Gln Thr Met
 1   5   10   15
Thr Asp Arg Lys Leu Leu Phe Ile Val Ala Thr Val Ile Leu Ala Val
            20   25   30
Leu Val Ser Phe Ser Ala Glu Leu Glu Ala Arg Val Lys Pro Thr Asn
        35   40   45
Ala Val Met Leu Gln Gly Phe Gly Trp Asn Ser Gln Ser Arg Gly Thr
    50   55   60
Pro Ser Lys Trp Tyr Asn Leu Ile Gly Asp Arg Ala Lys Asp Ile Ala
65   70   75   80
Asp Leu Gly Val Thr Met Val Trp Phe Pro Pro Val Ser Arg Ser Val
                85   90   95
Ser Pro Gln Gly Tyr Leu Pro Gly Asp Trp Tyr Asp Leu Gly Thr Glu
            100   105   110
Lys Asp Pro Thr Phe Tyr Gly His Lys Asp Ser Leu Leu Tyr Ala Leu
        115   120   125
Ser Glu Leu Lys Lys His Gly Ile Ser Pro Ile Ala Asp Ile Val Ile
    130   135   140
Asn His Arg Cys Ala Ser His Gln Asp Lys Asn Gly Ile Trp Asn Ile
145   150   155   160
Tyr His Phe Pro Ser Gly Lys Ala Arg Trp Glu Gln Trp Ala Ile Cys
                165   170   175
Arg Gly Gln Phe Gly Gly Gln Gly Asn Pro Asp Thr Gly Glu Ser Tyr
            180   185   190
His Ala Ala Pro Asp Val Asp His Thr Asn Val Arg Val Gln Arg Asp
        195   200   205
Ile Ile Glu Trp Met Asn Trp Leu Lys Ser Leu Gly Phe Glu Gly Trp
    210   215   220
Arg Tyr Asp Tyr Ser Lys Gly Tyr His Pro Arg Tyr Thr Ala Leu Tyr
225   230   235   240

```
Asp Arg Ser Thr Arg Pro Val Leu Ser Val Gly Glu Ile Trp Thr Asn
                245                 250                 255
Met Ser Phe Asn Gly Ser Tyr Leu Asn Pro Asn Gln Asp Ala His Arg
            260                 265                 270
Gln Gln Leu Cys Asp Trp Leu Asp Gly Ala Gly Asn Tyr Ala Thr Ala
        275                 280                 285
Phe Asp Phe Thr Thr Lys Gly Ile Leu Gln Val Ala Val Lys Gly Glu
    290                 295                 300
Tyr Trp Arg Leu Arg Asp Lys Asp Gly Lys Ala Ser Gly Leu Ile Gly
305                 310                 315                 320
Trp Trp Pro Glu Arg Ala Val Thr Phe Val Asp Asn His Asp Thr Gly
                325                 330                 335
Ser Gln Gln Ser His Trp Pro Phe Pro Gly Asn Lys Val Met Gln Gly
            340                 345                 350
Tyr Ala Tyr Ile Leu Thr His Pro Gly Ile Pro Cys Ile Phe Trp Glu
        355                 360                 365
His Val Tyr Asp Trp Lys Leu Arg Glu Pro Ile Lys Arg Leu Val Gln
    370                 375                 380
Ile Arg Lys Lys Tyr Gly Leu His Ser Lys Ser Ser Leu Lys Ile Val
385                 390                 395                 400
Arg Ala Glu Gln Asn Leu Tyr Ala Ala Ile Ile Asp Gly Lys Val Ala
                405                 410                 415
Leu Lys Leu Gly Ser Ala Asp Trp Ser Pro Pro Gly Asn Asp Phe Glu
            420                 425                 430
Leu Leu Thr His Gly Asp Gln Phe Ala Val Trp Gly Arg Lys
        435                 440                 445
```

<210> 456
<211> 1734
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 456

```
gtgatgcaac tgaatccctg gttttctgcc tctttaaaag ccgcaggcct tgcgaccgct      60
ctggcagcgg tcagcgcttg ccagccggca cccgagtcgg gttatcaagc cccaacggca     120
gcgcctgtcg agcagaccga agcgccctgg tgggacgatg cggttttcta ccaaatctgg     180
ccgcgtagtt tttacgatac cagcggtgat ggtcatggcg actttaatgg catgaccgcc     240
aaactgcctt atttgtcgga gcttggcatt gatgccatct ggctgacgcc aatgtttgag     300
gcgccgtctt atcatggcta tgactttacc gagttttatc aggtggaatc ggactatggc     360
accatggccg acttcgagcg ctttttgcaa caggctgaag agcgcggtat ccgggtgatt     420
ctggatttgg tgatcaacca catttcccgt gaacaccagt ggtttcagcg ctccgccgag     480
cgcattgagc cgtacagcga ttactttatc tggcgggaaa gcctgccaga aaccggcaaa     540
ggctggggtc atgcctggtc agataacgat atgcctgaag cggtctggca ttttgatgag     600
acccggggcg agtactacta cgccgccttt ggtgccaccc agccggatct gaatttggag     660
cacccggatg tgatcgaaga aatgaaacgg atggccaagt tctggctgga taaggggtc      720
gatggcttcc ggctcgatgc ggtgcgcttt gtgatagagc gaggcccgaa aggccaggcc     780
gacttgcccg agaccattga ttactggaaa gacttcaatg cctttgtcaa aagcgtcaat     840
ccagaagcct acttagtggg cgaagcctgg gtcgatattc cggtggcggc gaaatactgg     900
ggggaaggca aagggctgga tcagggcttt gattttgaat ttggctacat cgtgctggat     960
atgctgcagc aaggcttggc cacagaagcc cagtttggca ccatgagcca gcagtccggt    1020
caggcggaag acggcgcctt gcatccgctg gcgcagaatg tgcagcgccg cttggagtcg    1080
ggcgcaccgc tggccttctt cacgcccttt ttaaccaatc acgatcagga cagagttggc    1140
tttcaattgc tggaaaacca ggccaaagcc aaacaagcgg cggcactgtt attcagctcg    1200
ccgggcagca agtacatcta ctacggtgaa gaaattggca tgactcagct caaaagtggc    1260
catgatgtct acaagcgggc gccgatgttg tgggatcaag ggcatcaggc cggctttacc    1320
caggcagaaa aggcctgggt ggagcaaaaa gatttattca gccgccattt tgatgcctgg    1380
tggccggatt tcctgacaca gcagctggca gtcggcgatc ggacagtggc agcacagcgc    1440
gctcaaccgg actccttgtg gaacttgcac cgtctgctga tcgcgatgca aaagcagcgg    1500
ccggaaatgc gtttggcggg ctcgtatcag gtgcaatcgc cacagcaggg tattttgacg    1560
gtggtgcggg agctggataa cagccgctcg gtctttgtac tgaatctggc tggtgaaacg    1620
gccgatatcg catcggtgat ggctgcctat ggcgagctgt cggtcacttg gggctttgac    1680
ctggacggca gcttgctggc gaccgatggt ctggtcatct tgcaatcgcg gtaa           1734
```

<210> 457
<211> 577

<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(26)

<221> DOMAIN
<222> (58)...(440)
<223> Alpha amylase, catalytic domain

<400> 457

```
Met Met Gln Leu Asn Pro Trp Phe Ser Ala Ser Leu Lys Ala Ala Gly
 1               5                  10                  15
Leu Ala Thr Ala Leu Ala Ala Val Ser Ala Cys Gln Pro Ala Pro Glu
            20                  25                  30
Ser Gly Tyr Gln Ala Pro Thr Ala Ala Pro Val Glu Gln Thr Glu Ala
        35                  40                  45
Pro Trp Trp Asp Asp Ala Val Phe Tyr Gln Ile Trp Pro Arg Ser Phe
    50                  55                  60
Tyr Asp Thr Ser Gly Asp Gly His Gly Asp Phe Asn Gly Met Thr Ala
65                  70                  75                  80
Lys Leu Pro Tyr Leu Ser Glu Leu Gly Ile Asp Ala Ile Trp Leu Thr
                85                  90                  95
Pro Met Phe Glu Ala Pro Ser Tyr His Gly Tyr Asp Phe Thr Glu Phe
            100                 105                 110
Tyr Gln Val Glu Ser Asp Tyr Gly Thr Met Ala Asp Phe Glu Arg Phe
        115                 120                 125
Leu Gln Gln Ala Glu Glu Arg Gly Ile Arg Val Ile Leu Asp Leu Val
    130                 135                 140
Ile Asn His Ile Ser Arg Glu His Gln Trp Phe Gln Arg Ser Ala Glu
145                 150                 155                 160
Arg Ile Glu Pro Tyr Ser Asp Tyr Phe Ile Trp Arg Glu Ser Leu Pro
                165                 170                 175
Glu Thr Gly Lys Gly Trp Gly His Ala Trp Ser Asp Asn Asp Met Pro
            180                 185                 190
Glu Ala Val Trp His Phe Asp Glu Thr Arg Gly Glu Tyr Tyr Tyr Ala
        195                 200                 205
Ala Phe Gly Ala Thr Gln Pro Asp Leu Asn Leu Glu His Pro Asp Val
    210                 215                 220
Ile Glu Glu Met Lys Arg Met Ala Lys Phe Trp Leu Asp Lys Gly Val
225                 230                 235                 240
Asp Gly Phe Arg Leu Asp Ala Val Arg Phe Val Ile Glu Arg Gly Pro
                245                 250                 255
Lys Gly Gln Ala Asp Leu Pro Glu Thr Ile Asp Tyr Trp Lys Asp Phe
            260                 265                 270
Asn Ala Phe Val Lys Ser Val Asn Pro Glu Ala Tyr Leu Val Gly Glu
        275                 280                 285
Ala Trp Val Asp Ile Pro Val Ala Ala Lys Tyr Trp Gly Glu Gly Lys
    290                 295                 300
Gly Leu Asp Gln Gly Phe Asp Phe Glu Phe Gly Tyr Ile Val Leu Asp
305                 310                 315                 320
Met Leu Gln Gln Gly Leu Ala Thr Glu Ala Gln Phe Gly Thr Met Ser
                325                 330                 335
Gln Gln Ser Gly Gln Ala Glu Asp Gly Ala Leu His Pro Leu Ala Gln
            340                 345                 350
Asn Val Gln Arg Arg Leu Glu Ser Gly Ala Pro Leu Ala Phe Phe Thr
        355                 360                 365
Pro Phe Leu Thr Asn His Asp Gln Asp Arg Val Gly Phe Gln Leu Leu
    370                 375                 380
Glu Asn Gln Ala Lys Ala Lys Gln Ala Ala Ala Leu Leu Phe Ser Ser
385                 390                 395                 400
Pro Gly Ser Lys Tyr Ile Tyr Tyr Gly Glu Glu Ile Gly Met Thr Gln
                405                 410                 415
Leu Lys Ser Gly His Asp Val Tyr Lys Arg Ala Pro Met Leu Trp Asp
            420                 425                 430
```

Gln Gly His Gln Ala Gly Phe Thr Gln Ala Glu Lys Ala Trp Val Glu
435 440 445
Gln Lys Asp Leu Phe Ser Arg His Phe Asp Ala Trp Trp Pro Asp Phe
450 455 460
Leu Thr Gln Gln Leu Ala Val Gly Asp Arg Thr Val Ala Ala Gln Arg
465 470 475 480
Ala Gln Pro Asp Ser Leu Trp Asn Leu His Arg Leu Leu Ile Ala Met
485 490 495
Gln Lys Gln Arg Pro Glu Met Arg Leu Ala Gly Ser Tyr Gln Val Gln
500 505 510
Ser Pro Gln Gln Gly Ile Leu Thr Val Val Arg Glu Leu Asp Asn Ser
515 520 525
Arg Ser Val Phe Val Leu Asn Leu Ala Gly Glu Thr Ala Asp Ile Ala
530 535 540
Ser Val Met Ala Ala Tyr Gly Glu Leu Ser Val Thr Trp Gly Phe Asp
545 550 555 560
Leu Asp Gly Ser Leu Leu Ala Thr Asp Gly Leu Val Ile Leu Gln Ser
565 570 575
Arg

<210> 458
<211> 1698
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 458
atgttcaagg tcagtttaag gagtaaagac atgaagaagt taagtttaat agttaccatt 60
ttggtgcttg ctttgacctt atcagcgtgt caaagagaga ccgaagaaga agttggcgaa 120
acctatgact ttttgagttt gtttccgaat catcgaaact attatcaaat actcgttcgt 180
agctttgccg acggcaatgg cgatggcatc ggtgatttta ttggcttaaa agaagccctt 240
ccacatcttg ccgatttagg gattgaaggc ttatggctgt tgccgattca tccttcacca 300
agttaccacg ggtatgacgt gacagacttt agagcgatca atccggatta tggtacgatg 360
gaagattttg aagcgtttat tgcagaagca cagcgtttga acattgacgt cgtgatcgat 420
tatgtcatca atcattcttc caatcaacac ccttggtttc aaaatttcaa acagggtatc 480
gccccttacg atcaatttta tcgtcgtatt gataacactg atcccagatg gaacacgcgt 540
ggttcttggg gtcaaagcat ctggcactca ctgggggatg gaaccgctta tgtgggctat 600
tttggcgggt acatgcctga tttgaactgg tcgaatgacg ctttagttga agaaatgatc 660
gacatcggtc ttttctggct tgaaaaaggc gtgagtgggt tccgtttaga tgcggccttg 720
catttgttag ccaataacga agtgccgtca agcaccccag cgttagaaga aaccttattt 780
atgctagagt attttgagtt tcgtttaaaa gagaaatttc ctgatacttt cattgtgggt 840
gaaatttggg atgcttttag tatttacaac atgttttacc gtagtatgga ctcagttttt 900
cactttgact ttggtcactt ggtcgtgaac acgattaaca ctggttctaa tcgcaactat 960
gctgatcaaa tcgtccgctg gcacaatcaa gcgcaagcag tagacgatgt ggttatcgaa 1020
gcgccgtttt taagaaatca tgatcaaaac cgtctggctt cagcctcctc tccaggctca 1080
ccgaatctcg gccccaacgc ccatcctgat aaactgcgtt tagccgcaga gatgttattg 1140
actgtccctg gaagtccgtt tatttattat ggtgaagaac tcggcatgaa aggggaagcc 1200
tctggtcgag ccccaacgtg ggattcaacg attcgtttac cttacttgtg ggacgacgaa 1260
cgtcaaccaa cgtggacgtt tgaccggttt gggtttgtcg atacgtttaa tgtcgatgtc 1320
ccgaccttgc cagaacaaca agccgatccg ttatcactct ttaacacata ccgtacctta 1380
ctacatttaa gacaagcgca tcctgcctta agactcggta ctattgaagc gtttgaagaa 1440
aatcatgcgg gcttacaagg gttttaccgt atttttgaat acagcacctt ccaagaagtg 1500
gtcttagtgc ttcataacgt gagtgatgag actttgaacg tgactcgttt acctgaaggt 1560
gagatgattt acctaagtgc acatcttgaa gattttccag cgctcgaaca actcaatcag 1620
acgcttgata tcggccctag aagtacgctc atgattcgtt taagtgagac agcgatggag 1680
cgtttcttca atgaatga 1698

<210> 459
<211> 565
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> DOMAIN
<222> (56)...(464)
<223> Alpha amylase, catalytic domain

<221> SIGNAL
<222> (1)...(29)

<400> 459

```
Met Phe Lys Val Ser Leu Arg Ser Lys Asp Met Lys Lys Leu Ser Leu
 1               5                  10                  15
Ile Val Thr Ile Leu Val Leu Ala Leu Thr Leu Ser Ala Cys Gln Arg
            20                  25                  30
Glu Thr Glu Glu Glu Val Gly Glu Thr Tyr Asp Phe Leu Ser Leu Phe
        35                  40                  45
Pro Asn His Arg Asn Tyr Tyr Gln Ile Leu Val Arg Ser Phe Ala Asp
    50                  55                  60
Gly Asn Gly Asp Gly Ile Gly Asp Phe Ile Gly Leu Lys Glu Ala Leu
65                  70                  75                  80
Pro His Leu Ala Asp Leu Gly Ile Glu Gly Leu Trp Leu Leu Pro Ile
                85                  90                  95
His Pro Ser Pro Ser Tyr His Gly Tyr Asp Val Thr Asp Phe Arg Ala
            100                 105                 110
Ile Asn Pro Asp Tyr Gly Thr Met Glu Asp Phe Glu Ala Phe Ile Ala
        115                 120                 125
Glu Ala Gln Arg Leu Asn Ile Asp Val Val Ile Asp Tyr Val Ile Asn
    130                 135                 140
His Ser Ser Asn Gln His Pro Trp Phe Gln Asn Phe Lys Gln Gly Ile
145                 150                 155                 160
Ala Pro Tyr Asp Gln Phe Tyr Arg Arg Ile Asp Asn Thr Asp Pro Arg
                165                 170                 175
Trp Asn Thr Arg Gly Ser Trp Gly Gln Ser Ile Trp His Ser Leu Gly
            180                 185                 190
Asp Gly Thr Ala Tyr Val Gly Tyr Phe Gly Gly Tyr Met Pro Asp Leu
        195                 200                 205
Asn Trp Ser Asn Asp Ala Leu Val Glu Glu Met Ile Asp Ile Gly Leu
    210                 215                 220
Phe Trp Leu Glu Lys Gly Val Ser Gly Phe Arg Leu Asp Ala Ala Leu
225                 230                 235                 240
His Leu Leu Ala Asn Asn Glu Val Pro Ser Ser Thr Pro Ala Leu Glu
                245                 250                 255
Glu Thr Leu Phe Met Leu Glu Tyr Phe Glu Phe Arg Leu Lys Glu Lys
            260                 265                 270
Phe Pro Asp Thr Phe Ile Val Gly Glu Ile Trp Asp Ala Phe Ser Ile
        275                 280                 285
Tyr Asn Met Phe Tyr Arg Ser Met Asp Ser Val Phe His Phe Asp Phe
    290                 295                 300
Gly His Leu Val Val Asn Thr Ile Asn Thr Gly Ser Asn Arg Asn Tyr
305                 310                 315                 320
Ala Asp Gln Ile Val Arg Trp His Asn Gln Ala Gln Ala Val Asp Asp
                325                 330                 335
Val Val Ile Glu Ala Pro Phe Leu Arg Asn His Asp Gln Asn Arg Leu
            340                 345                 350
Ala Ser Ala Ser Ser Pro Gly Ser Pro Asn Leu Gly Pro Asn Ala His
        355                 360                 365
Pro Asp Lys Leu Arg Leu Ala Ala Glu Met Leu Leu Thr Val Pro Gly
    370                 375                 380
Ser Pro Phe Ile Tyr Tyr Gly Glu Glu Leu Gly Met Lys Gly Glu Ala
385                 390                 395                 400
Ser Gly Arg Ala Pro Thr Trp Asp Ser Thr Ile Arg Leu Pro Tyr Leu
                405                 410                 415
Trp Asp Asp Glu Arg Gln Pro Thr Trp Thr Phe Asp Arg Phe Gly Phe
            420                 425                 430
Val Asp Thr Phe Asn Val Asp Val Pro Thr Leu Pro Glu Gln Gln Ala
        435                 440                 445
Asp Pro Leu Ser Leu Phe Asn Thr Tyr Arg Thr Leu Leu His Leu Arg
    450                 455                 460
Gln Ala His Pro Ala Leu Arg Leu Gly Thr Ile Glu Ala Phe Glu Glu
465                 470                 475                 480
```

```
Asn His Ala Gly Leu Gln Gly Phe Tyr Arg Ile Phe Glu Tyr Ser Thr
                485                 490                 495
Phe Gln Glu Val Val Leu Val Leu His Asn Val Ser Asp Glu Thr Leu
            500                 505                 510
Asn Val Thr Arg Leu Pro Glu Gly Glu Met Ile Tyr Leu Ser Ala His
        515                 520                 525
Leu Glu Asp Phe Pro Ala Leu Glu Gln Leu Asn Gln Thr Leu Asp Ile
    530                 535                 540
Gly Pro Arg Ser Thr Leu Met Ile Arg Leu Ser Glu Thr Ala Met Glu
545                 550                 555                 560
Arg Phe Phe Asn Glu
                565

<210> 460
<211> 1752
<212> DNA
<213> Cochliobolus heterostrophus ATCC 48331

<220>
<221> CDS
<222> (1)...(1752)
<223> Exon

<400> 460
atg cca caa gat tta cga cct tgg tgg caa act gcc atc ggt tat caa       48
Met Pro Gln Asp Leu Arg Pro Trp Trp Gln Thr Ala Ile Gly Tyr Gln
 1               5                  10                  15

ata tgg ccg gcg tct ttc aaa gat tcg aac agt gat ggc ata ggc gat       96
Ile Trp Pro Ala Ser Phe Lys Asp Ser Asn Ser Asp Gly Ile Gly Asp
             20                  25                  30

ctc aat gga gtt ctc agc aag ctc gac tac ctc cga tca ctg gaa gtc      144
Leu Asn Gly Val Leu Ser Lys Leu Asp Tyr Leu Arg Ser Leu Glu Val
         35                  40                  45

gat ctc atc tgg ttt tct cct gtc tat gat tcg ccc aac aga gac atg      192
Asp Leu Ile Trp Phe Ser Pro Val Tyr Asp Ser Pro Asn Arg Asp Met
     50                  55                  60

gga tat gat gtc tct gat tac gaa tcc att cac cca cag tat ggc acc      240
Gly Tyr Asp Val Ser Asp Tyr Glu Ser Ile His Pro Gln Tyr Gly Thr
 65                  70                  75                  80

att gcc gac atg gaa aac ctc atc aaa gag gcc aag aaa cgt aac atc      288
Ile Ala Asp Met Glu Asn Leu Ile Lys Glu Ala Lys Lys Arg Asn Ile
                 85                  90                  95

aaa att gtc atg gat atg gtc aca aat cat tcg tct gac caa cac cgc      336
Lys Ile Val Met Asp Met Val Thr Asn His Ser Ser Asp Gln His Arg
            100                 105                 110

tgg ttc ttg gag agt tgc aaa tct cgg tca ggg cca tat agt gac ttt      384
Trp Phe Leu Glu Ser Cys Lys Ser Arg Ser Gly Pro Tyr Ser Asp Phe
        115                 120                 125

tac atc tgg cgg gat cca aaa tac gat agc gaa ggt aac cga caa cca      432
Tyr Ile Trp Arg Asp Pro Lys Tyr Asp Ser Glu Gly Asn Arg Gln Pro
    130                 135                 140

cca aat aat tgg cgc act ggg tca cga tat gga aaa agt aca tgg aag      480
Pro Asn Asn Trp Arg Thr Gly Ser Arg Tyr Gly Lys Ser Thr Trp Lys
145                 150                 155                 160

tac gtc gaa gca aga gat cag tac tat ttc tgc ttt gga act act tat      528
Tyr Val Glu Ala Arg Asp Gln Tyr Tyr Phe Cys Phe Gly Thr Thr Tyr
                165                 170                 175
```

```
cag cct aac ttg aat tgg ttc tgc gac gat ttg aga cgc gca ata tac      576
Gln Pro Asn Leu Asn Trp Phe Cys Asp Asp Leu Arg Arg Ala Ile Tyr
            180                 185                 190

gac agc gcg gtg aaa ttt tgg ttg gac aga ggt gtc gat ggg ttc agg      624
Asp Ser Ala Val Lys Phe Trp Leu Asp Arg Gly Val Asp Gly Phe Arg
        195                 200                 205

atg gat ttg gtc ggg tta tac tgg aaa gat cca acc ttt cct gat gca      672
Met Asp Leu Val Gly Leu Tyr Trp Lys Asp Pro Thr Phe Pro Asp Ala
    210                 215                 220

aaa gaa gtt tat cct gga gag gcg ctg caa ccg ctt gac gga tcg cgg      720
Lys Glu Val Tyr Pro Gly Glu Ala Leu Gln Pro Leu Asp Gly Ser Arg
225                 230                 235                 240

tgt ttg aat ggc gaa gaa gtg cat gtc tgg tta aag gag ttg aag gaa      768
Cys Leu Asn Gly Glu Glu Val His Val Trp Leu Lys Glu Leu Lys Glu
                245                 250                 255

aga atc acg caa gac ttt gga cag gac atc gtg ttg att ggg gag ctc      816
Arg Ile Thr Gln Asp Phe Gly Gln Asp Ile Val Leu Ile Gly Glu Leu
            260                 265                 270

ccc tcg act gac cgg gag aca ttc cta cga tac ata tct cct gaa tcg      864
Pro Ser Thr Asp Arg Glu Thr Phe Leu Arg Tyr Ile Ser Pro Glu Ser
        275                 280                 285

aga gaa cta gag atg gct ctc gat aca gac atc ttc atc gtg gga aat      912
Arg Glu Leu Glu Met Ala Leu Asp Thr Asp Ile Phe Ile Val Gly Asn
    290                 295                 300

gaa tgg aca gat ggt ctg tac gag ctc aaa aag cca gag cta ccg cta      960
Glu Trp Thr Asp Gly Leu Tyr Glu Leu Lys Lys Pro Glu Leu Pro Leu
305                 310                 315                 320

ttc aaa gat gcc gtc ctg aag acg cag agc ctt ctt gac gat gga ggc     1008
Phe Lys Asp Ala Val Leu Lys Thr Gln Ser Leu Leu Asp Asp Gly Gly
                325                 330                 335

tgg cct act gca ttc ctc gag aat cat gat ttc gct cgc agt gta tcc     1056
Trp Pro Thr Ala Phe Leu Glu Asn His Asp Phe Ala Arg Ser Val Ser
            340                 345                 350

cgt ttt ggg cca ggt gag ggg acc tat cgc gaa gct gcc gca agg atg     1104
Arg Phe Gly Pro Gly Glu Gly Thr Tyr Arg Glu Ala Ala Ala Arg Met
        355                 360                 365

ctt gct ctc atg gct gca aca ctt tcc gga acc ttt ttc atc tac cag     1152
Leu Ala Leu Met Ala Ala Thr Leu Ser Gly Thr Phe Phe Ile Tyr Gln
    370                 375                 380

ggt caa gaa att gga atg acc aac gtc cca tca cat tgg agg cga gat     1200
Gly Gln Glu Ile Gly Met Thr Asn Val Pro Ser His Trp Arg Arg Asp
385                 390                 395                 400

gac ttc aaa gat aac gcc gac ctc gaa cac atc gaa gat att gat cca     1248
Asp Phe Lys Asp Asn Ala Asp Leu Glu His Ile Glu Asp Ile Asp Pro
                405                 410                 415

atc aaa gac gct gag gcg gtg gag aaa tcc gtt gat gcg ctt acc aca     1296
Ile Lys Asp Ala Glu Ala Val Glu Lys Ser Val Asp Ala Leu Thr Thr
            420                 425                 430

tgg ggg cga gat aat gga aga acg cca gtc cag tgg tca gac aaa gat     1344
Trp Gly Arg Asp Asn Gly Arg Thr Pro Val Gln Trp Ser Asp Lys Asp
        435                 440                 445
```

```
cat gcc ggg ttt agc gac gtg aaa ccc tgg att agg gtc atc gac aac     1392
His Ala Gly Phe Ser Asp Val Lys Pro Trp Ile Arg Val Ile Asp Asn
    450                 455                 460

tac act tac ata aac gtc gaa gag caa ctt gga aac ccg aac agc ata     1440
Tyr Thr Tyr Ile Asn Val Glu Glu Gln Leu Gly Asn Pro Asn Ser Ile
465                 470                 475                 480

cgt tct ttc tgg gta aag atg atg aag atg agg aaa gag atg tat gat     1488
Arg Ser Phe Trp Val Lys Met Met Lys Met Arg Lys Glu Met Tyr Asp
                485                 490                 495

gtg tta gcc tgc ggg cat ttt cgt ctc gtt gac agg gac aac ctc agt     1536
Val Leu Ala Cys Gly His Phe Arg Leu Val Asp Arg Asp Asn Leu Ser
            500                 505                 510

gtg ttc tcg tat acg aaa aca agt gct gat gga aag acc gag atg cta     1584
Val Phe Ser Tyr Thr Lys Thr Ser Ala Asp Gly Lys Thr Glu Met Leu
        515                 520                 525

gtg gta ttg aac ttt tcg agc gag gag gct cgg gct ccg agg caa atc     1632
Val Val Leu Asn Phe Ser Ser Glu Glu Ala Arg Ala Pro Arg Gln Ile
    530                 535                 540

ggg aaa gat cta atg cat ctc gtt tgc acg aca gct gcg atg tcg acg     1680
Gly Lys Asp Leu Met His Leu Val Cys Thr Thr Ala Ala Met Ser Thr
545                 550                 555                 560

aag act agc tta gat gaa gtc aca ttg cag cca tgg gag ggt cgc ttg     1728
Lys Thr Ser Leu Asp Glu Val Thr Leu Gln Pro Trp Glu Gly Arg Leu
                565                 570                 575

tat cga ctg gca cca aag ttg tag                                    1752
Tyr Arg Leu Ala Pro Lys Leu  *
            580
```

<210> 461
<211> 583
<212> PRT
<213> Cochliobolus heterostrophus ATCC 48331

<220>
<221> DOMAIN
<222> (16)...(421)

<400> 461

```
Met Pro Gln Asp Leu Arg Pro Trp Trp Gln Thr Ala Ile Gly Tyr Gln
 1               5                  10                  15
Ile Trp Pro Ala Ser Phe Lys Asp Ser Asn Ser Asp Gly Ile Gly Asp
            20                  25                  30
Leu Asn Gly Val Leu Ser Lys Leu Asp Tyr Leu Arg Ser Leu Glu Val
        35                  40                  45
Asp Leu Ile Trp Phe Ser Pro Val Tyr Asp Ser Pro Asn Arg Asp Met
    50                  55                  60
Gly Tyr Asp Val Ser Asp Tyr Glu Ser Ile His Pro Gln Tyr Gly Thr
65                  70                  75                  80
Ile Ala Asp Met Glu Asn Leu Ile Lys Glu Ala Lys Lys Arg Asn Ile
                85                  90                  95
Lys Ile Val Met Asp Met Val Thr Asn His Ser Ser Asp Gln His Arg
            100                 105                 110
Trp Phe Leu Glu Ser Cys Lys Ser Arg Ser Gly Pro Tyr Ser Asp Phe
        115                 120                 125
Tyr Ile Trp Arg Asp Pro Lys Tyr Asp Ser Glu Gly Asn Arg Gln Pro
    130                 135                 140
Pro Asn Asn Trp Arg Thr Gly Ser Arg Tyr Gly Lys Ser Thr Trp Lys
145                 150                 155                 160
Tyr Val Glu Ala Arg Asp Gln Tyr Tyr Phe Cys Phe Gly Thr Thr Tyr
```

```
                165                 170                 175
Gln Pro Asn Leu Asn Trp Phe Cys Asp Asp Leu Arg Arg Ala Ile Tyr
            180                 185                 190
Asp Ser Ala Val Lys Phe Trp Leu Asp Arg Gly Val Asp Gly Phe Arg
        195                 200                 205
Met Asp Leu Val Gly Leu Tyr Trp Lys Asp Pro Thr Phe Pro Asp Ala
    210                 215                 220
Lys Glu Val Tyr Pro Gly Glu Ala Leu Gln Pro Leu Asp Gly Ser Arg
225                 230                 235                 240
Cys Leu Asn Gly Glu Glu Val His Val Trp Leu Lys Glu Leu Lys Glu
                245                 250                 255
Arg Ile Thr Gln Asp Phe Gly Gln Asp Ile Val Leu Ile Gly Glu Leu
            260                 265                 270
Pro Ser Thr Asp Arg Glu Thr Phe Leu Arg Tyr Ile Ser Pro Glu Ser
        275                 280                 285
Arg Glu Leu Glu Met Ala Leu Asp Thr Asp Ile Phe Ile Val Gly Asn
    290                 295                 300
Glu Trp Thr Asp Gly Leu Tyr Glu Leu Lys Lys Pro Glu Leu Pro Leu
305                 310                 315                 320
Phe Lys Asp Ala Val Leu Lys Thr Gln Ser Leu Leu Asp Asp Gly Gly
                325                 330                 335
Trp Pro Thr Ala Phe Leu Glu Asn His Asp Phe Ala Arg Ser Val Ser
            340                 345                 350
Arg Phe Gly Pro Gly Glu Gly Thr Tyr Arg Glu Ala Ala Ala Arg Met
        355                 360                 365
Leu Ala Leu Met Ala Ala Thr Leu Ser Gly Thr Phe Phe Ile Tyr Gln
    370                 375                 380
Gly Gln Glu Ile Gly Met Thr Asn Val Pro Ser His Trp Arg Arg Asp
385                 390                 395                 400
Asp Phe Lys Asp Asn Ala Asp Leu Glu His Ile Glu Asp Ile Asp Pro
                405                 410                 415
Ile Lys Asp Ala Glu Ala Val Glu Lys Ser Val Asp Ala Leu Thr Thr
            420                 425                 430
Trp Gly Arg Asp Asn Gly Arg Thr Pro Val Gln Trp Ser Asp Lys Asp
        435                 440                 445
His Ala Gly Phe Ser Asp Val Lys Pro Trp Ile Arg Val Ile Asp Asn
    450                 455                 460
Tyr Thr Tyr Ile Asn Val Glu Glu Gln Leu Gly Asn Pro Asn Ser Ile
465                 470                 475                 480
Arg Ser Phe Trp Val Lys Met Met Lys Met Arg Lys Glu Met Tyr Asp
                485                 490                 495
Val Leu Ala Cys Gly His Phe Arg Leu Val Asp Arg Asp Asn Leu Ser
            500                 505                 510
Val Phe Ser Tyr Thr Lys Thr Ser Ala Asp Gly Lys Thr Glu Met Leu
        515                 520                 525
Val Val Leu Asn Phe Ser Ser Glu Glu Ala Arg Ala Pro Arg Gln Ile
    530                 535                 540
Gly Lys Asp Leu Met His Leu Val Cys Thr Thr Ala Ala Met Ser Thr
545                 550                 555                 560
Lys Thr Ser Leu Asp Glu Val Thr Leu Gln Pro Trp Glu Gly Arg Leu
                565                 570                 575
Tyr Arg Leu Ala Pro Lys Leu
            580

<210> 462
<211> 1252
<212> DNA
<213> Fungi

<220>
<221> CDS
<222> (1)...(114)
<223> Exon

<221> CDS
<222> (161)...(1252)
<223> Exon
```

<400> 462

```
atg tca aga tct tct acc att ctt ttt gtg ttg gca gca gcc aat ttg       48
Met Ser Arg Ser Ser Thr Ile Leu Phe Val Leu Ala Ala Ala Asn Leu
 1               5                  10                  15

gcg tct cta gta gat gct cat ggc tac atg aac ata ccc tcc agt cga       96
Ala Ser Leu Val Asp Ala His Gly Tyr Met Asn Ile Pro Ser Ser Arg
            20                  25                  30

aca agc atc ggc cat tca gtaagattcc ataaatccat ttctattacc             144
Thr Ser Ile Gly His Ser
         35

acattaacaa ttgaag gtt ggc atc gac aca tgt ccc gaa tgt act att ctt    196
                  Val Gly Ile Asp Thr Cys Pro Glu Cys Thr Ile Leu
                      40                  45                  50

gag ccc gtc tct tca tgg cct gat ctc gac gtc gct cca gtt gga cgc      244
Glu Pro Val Ser Ser Trp Pro Asp Leu Asp Val Ala Pro Val Gly Arg
                55                  60                  65

agc ggt cct tgc ggc tac aat gct cgg gtc agc gtc gac tac aat tac      292
Ser Gly Pro Cys Gly Tyr Asn Ala Arg Val Ser Val Asp Tyr Asn Tyr
            70                  75                  80

cct tcc aca aac tgg ggt aac acc ccc gtc acc acc tat aca gct ggt      340
Pro Ser Thr Asn Trp Gly Asn Thr Pro Val Thr Thr Tyr Thr Ala Gly
         85                  90                  95

caa aca gtc gat gta cag tgg tgc gta gac aac aac gga gac cat ggc      388
Gln Thr Val Asp Val Gln Trp Cys Val Asp Asn Asn Gly Asp His Gly
    100                 105                 110

ggt atg ttc agc tgg cgc atc tgt caa gac caa tcc atc gtg gat aag      436
Gly Met Phe Ser Trp Arg Ile Cys Gln Asp Gln Ser Ile Val Asp Lys
115                 120                 125                 130

ctc atc aca ccc gga tac atc ccc acg gat gcc gag aaa caa gcc gct      484
Leu Ile Thr Pro Gly Tyr Ile Pro Thr Asp Ala Glu Lys Gln Ala Ala
                135                 140                 145

gag gat tgt ttc gac gcc gga tta ctc ccc tgt acc gac gtc acc gga      532
Glu Asp Cys Phe Asp Ala Gly Leu Leu Pro Cys Thr Asp Val Thr Gly
            150                 155                 160

gcc acg tgc gat tac aac gca gat tgc act ccc gga caa gcc tgt tac      580
Ala Thr Cys Asp Tyr Asn Ala Asp Cys Thr Pro Gly Gln Ala Cys Tyr
        165                 170                 175

cga aac gac tgg ttt act tgc aag gga ttc aac agc ggc aca aga tgt      628
Arg Asn Asp Trp Phe Thr Cys Lys Gly Phe Asn Ser Gly Thr Arg Cys
    180                 185                 190

caa gga gtg gat ggc gcc gct ctc ggc tct tgc tac act tcc atc gca      676
Gln Gly Val Asp Gly Ala Ala Leu Gly Ser Cys Tyr Thr Ser Ile Ala
195                 200                 205                 210

ggt gga tat aca gtg acc aag aaa gtc aag atc cca aac tac gca tca      724
Gly Gly Tyr Thr Val Thr Lys Lys Val Lys Ile Pro Asn Tyr Ala Ser
                215                 220                 225

aat cac acc ctc ata tcc ctg aaa tgg aat tcc ttc caa aca ggc caa      772
Asn His Thr Leu Ile Ser Leu Lys Trp Asn Ser Phe Gln Thr Gly Gln
            230                 235                 240

atc tat cta acc tgc gcc gat atc gcc att gtc ggt tcc ggc agc ggc      820
Ile Tyr Leu Thr Cys Ala Asp Ile Ala Ile Val Gly Ser Gly Ser Gly
        245                 250                 255
```

```
ggc ggc tcc tca agt acc acg agc aaa aca tct tcc acc tcc acc acc      868
Gly Gly Ser Ser Ser Thr Thr Ser Lys Thr Ser Ser Thr Ser Thr Thr
    260                 265                 270

ctc agc aca tcg acc aaa tcg agc acc acc acc agc gcc acc acg gca      916
Leu Ser Thr Ser Thr Lys Ser Ser Thr Thr Thr Ser Ala Thr Thr Ala
275                 280                 285                 290

acc aca agc tgc gcc tca acc cct tct ctg att ccc gta acc ttc aac      964
Thr Thr Ser Cys Ala Ser Thr Pro Ser Leu Ile Pro Val Thr Phe Asn
                295                 300                 305

gag ctc gtc acc acc acc tac gga caa acc atc aaa gtc gcc ggc tcg     1012
Glu Leu Val Thr Thr Thr Tyr Gly Gln Thr Ile Lys Val Ala Gly Ser
            310                 315                 320

gtc gcc gct ctc gga aac tgg gat gtt tcc gcc gcc gtc gca ttg tca     1060
Val Ala Ala Leu Gly Asn Trp Asp Val Ser Ala Ala Val Ala Leu Ser
        325                 330                 335

gct gta aac tac acg agt gcg aat cca ttg tgg acc gga acc gtc atg     1108
Ala Val Asn Tyr Thr Ser Ala Asn Pro Leu Trp Thr Gly Thr Val Met
    340                 345                 350

ttg ccc ccc gga cag gtg gtg cag tat aag tat gtg aat gtg gcg agt     1156
Leu Pro Pro Gly Gln Val Val Gln Tyr Lys Tyr Val Asn Val Ala Ser
355                 360                 365                 370

agt ggg acg cca acg tgg gaa aaa gat ccc aat cgc gtt ttt acg gtg     1204
Ser Gly Thr Pro Thr Trp Glu Lys Asp Pro Asn Arg Val Phe Thr Val
                375                 380                 385

ccc gcg ggt tgt gcg aca aag ggg acc gtg ggg gat act tgg aga tag     1252
Pro Ala Gly Cys Ala Thr Lys Gly Thr Val Gly Asp Thr Trp Arg  *
            390                 395                 400
```

<210> 463
<211> 38
<212> PRT
<213> Fungi

<400> 463
```
Met Ser Arg Ser Ser Thr Ile Leu Phe Val Leu Ala Ala Ala Asn Leu
 1               5                  10                  15
Ala Ser Leu Val Asp Ala His Gly Tyr Met Asn Ile Pro Ser Ser Arg
            20                  25                  30
Thr Ser Ile Gly His Ser
        35
```

<210> 464
<211> 363
<212> PRT
<213> Fungi

<400> 464
```
Val Gly Ile Asp Thr Cys Pro Glu Cys Thr Ile Leu Glu Pro Val Ser
 1               5                  10                  15
Ser Trp Pro Asp Leu Asp Val Ala Pro Val Gly Arg Ser Gly Pro Cys
            20                  25                  30
Gly Tyr Asn Ala Arg Val Ser Val Asp Tyr Asn Tyr Pro Ser Thr Asn
        35                  40                  45
Trp Gly Asn Thr Pro Val Thr Thr Tyr Thr Ala Gly Gln Thr Val Asp
    50                  55                  60
Val Gln Trp Cys Val Asp Asn Asn Gly Asp His Gly Gly Met Phe Ser
65                  70                  75                  80
Trp Arg Ile Cys Gln Asp Gln Ser Ile Val Asp Lys Leu Ile Thr Pro
```

```
                85                  90                  95
Gly Tyr Ile Pro Thr Asp Ala Glu Lys Gln Ala Ala Glu Asp Cys Phe
            100                 105                 110
Asp Ala Gly Leu Leu Pro Cys Thr Asp Val Thr Gly Ala Thr Cys Asp
        115                 120                 125
Tyr Asn Ala Asp Cys Thr Pro Gly Gln Ala Cys Tyr Arg Asn Asp Trp
    130                 135                 140
Phe Thr Cys Lys Gly Phe Asn Ser Gly Thr Arg Cys Gln Gly Val Asp
145                 150                 155                 160
Gly Ala Ala Leu Gly Ser Cys Tyr Thr Ser Ile Ala Gly Gly Tyr Thr
                165                 170                 175
Val Thr Lys Lys Val Lys Ile Pro Asn Tyr Ala Ser Asn His Thr Leu
            180                 185                 190
Ile Ser Leu Lys Trp Asn Ser Phe Gln Thr Gly Gln Ile Tyr Leu Thr
        195                 200                 205
Cys Ala Asp Ile Ala Ile Val Gly Ser Gly Ser Gly Gly Gly Ser Ser
    210                 215                 220
Ser Thr Thr Ser Lys Thr Ser Ser Thr Ser Thr Thr Leu Ser Thr Ser
225                 230                 235                 240
Thr Lys Ser Ser Thr Thr Thr Ser Ala Thr Thr Ala Thr Thr Ser Cys
                245                 250                 255
Ala Ser Thr Pro Ser Leu Ile Pro Val Thr Phe Asn Glu Leu Val Thr
            260                 265                 270
Thr Thr Tyr Gly Gln Thr Ile Lys Val Ala Gly Ser Val Ala Ala Leu
        275                 280                 285
Gly Asn Trp Asp Val Ser Ala Ala Val Ala Leu Ser Ala Val Asn Tyr
    290                 295                 300
Thr Ser Ala Asn Pro Leu Trp Thr Gly Thr Val Met Leu Pro Pro Gly
305                 310                 315                 320
Gln Val Val Gln Tyr Lys Tyr Val Asn Val Ala Ser Ser Gly Thr Pro
                325                 330                 335
Thr Trp Glu Lys Asp Pro Asn Arg Val Phe Thr Val Pro Ala Gly Cys
            340                 345                 350
Ala Thr Lys Gly Thr Val Gly Asp Thr Trp Arg
        355                 360
```

<210> 465
<211> 1206
<212> DNA
<213> Fungi

<400> 465
atgtcaagat cttctaccat tctttttgtg ttggcagcag ccaatttggc gtctctagta 60
gatgctcatg gctacatgaa cataccctcc agtcgaacaa gcatcggcca ttcagttggc 120
atcgacacat gtcccgaatg tactattctt gagcccgtct cttcatggcc tgatctcgac 180
gtcgctccag ttggacgcag cggtccttgc ggctacaatg ctcgggtcag cgtcgactac 240
aattaccctt ccacaaactg gggtaacacc cccgtcacca cctatacagc tggtcaaaca 300
gtcgatgtac agtggtgcgt agacaacaac ggagaccatg gcggtatgtt cagctggcgc 360
atctgtcaag accaatccat cgtggataag ctcatcacac ccggatacat ccccacggat 420
gccgagaaac aagccgctga ggattgtttc gacgccggat tactcccctg taccgacgtc 480
accggagcca cgtgcgatta caacgcagat tgcactcccg gacaagcctg ttaccgaaac 540
gactggttta cttgcaaggg attcaacagc ggcacaagat gtcaaggagt ggatggcgcc 600
gctctcggct cttgctacac ttccatcgca ggtggatata cagtgaccaa gaaagtcaag 660
atcccaaact acgcatcaaa tcacaccctc atatccctga aatggaattc cttccaaaca 720
ggccaaatct atctaacctg cgccgatatc gccattgtcg gttccggcag cggcggcggc 780
tcctcaagta ccacgagcaa aacatcttcc acctccacca ccctcagcac atcgaccaaa 840
tcgagcacca ccaccagcgc caccacggca accacaagct gcgcctcaac cccttctctg 900
attcccgtaa ccttcaacga gctcgtcacc accacctacg gacaaaccat caaagtcgcc 960
ggctcggtcg ccgctctcgg aaactgggat gtttccgccg ccgtcgcatt gtcagctgta 1020
aactacacga gtgcgaatcc attgtggacc ggaaccgtca tgttgccccc cggacaggtg 1080
gtgcagtata agtatgtgaa tgtggcgagt agtgggacgc caacgtggga aaaagatccc 1140
aatcgcgttt ttacggtgcc cgcgggttgt gcgacaaagg ggaccgtggg ggatacttgg 1200
agatag 1206

<210> 466
<211> 401
<212> PRT
<213> Fungi

<220>
<221> SIGNAL
<222> (1)...(22)

<221> DOMAIN
<222> (300)...(396)
<223> Starch binding domain

<400> 466
Met Ser Arg Ser Ser Thr Ile Leu Phe Val Leu Ala Ala Ala Asn Leu
 1   5   10   15
Ala Ser Leu Val Asp Ala His Gly Tyr Met Asn Ile Pro Ser Ser Arg
            20   25   30
Thr Ser Ile Gly His Ser Val Gly Ile Asp Thr Cys Pro Glu Cys Thr
        35   40   45
Ile Leu Glu Pro Val Ser Ser Trp Pro Asp Leu Asp Val Ala Pro Val
    50   55   60
Gly Arg Ser Gly Pro Cys Gly Tyr Asn Ala Arg Val Ser Val Asp Tyr
65   70   75   80
Asn Tyr Pro Ser Thr Asn Trp Gly Asn Thr Pro Val Thr Thr Tyr Thr
                85   90   95
Ala Gly Gln Thr Val Asp Val Gln Trp Cys Val Asp Asn Asn Gly Asp
            100   105   110
His Gly Gly Met Phe Ser Trp Arg Ile Cys Gln Asp Gln Ser Ile Val
        115   120   125
Asp Lys Leu Ile Thr Pro Gly Tyr Ile Pro Thr Asp Ala Glu Lys Gln
    130   135   140
Ala Ala Glu Asp Cys Phe Asp Ala Gly Leu Leu Pro Cys Thr Asp Val
145   150   155   160
Thr Gly Ala Thr Cys Asp Tyr Asn Ala Asp Cys Thr Pro Gly Gln Ala
                165   170   175
Cys Tyr Arg Asn Asp Trp Phe Thr Cys Lys Gly Phe Asn Ser Gly Thr
            180   185   190
Arg Cys Gln Gly Val Asp Gly Ala Ala Leu Gly Ser Cys Tyr Thr Ser
        195   200   205
Ile Ala Gly Gly Tyr Thr Val Thr Lys Lys Val Lys Ile Pro Asn Tyr
    210   215   220
Ala Ser Asn His Thr Leu Ile Ser Leu Lys Trp Asn Ser Phe Gln Thr
225   230   235   240
Gly Gln Ile Tyr Leu Thr Cys Ala Asp Ile Ala Ile Val Gly Ser Gly
                245   250   255
Ser Gly Gly Gly Ser Ser Ser Thr Thr Ser Lys Thr Ser Ser Thr Ser
            260   265   270
Thr Thr Leu Ser Thr Ser Thr Lys Ser Ser Thr Thr Thr Ser Ala Thr
        275   280   285
Thr Ala Thr Thr Ser Cys Ala Ser Thr Pro Ser Leu Ile Pro Val Thr
    290   295   300
Phe Asn Glu Leu Val Thr Thr Thr Tyr Gly Gln Thr Ile Lys Val Ala
305   310   315   320
Gly Ser Val Ala Ala Leu Gly Asn Trp Asp Val Ser Ala Ala Val Ala
                325   330   335
Leu Ser Ala Val Asn Tyr Thr Ser Ala Asn Pro Leu Trp Thr Gly Thr
            340   345   350
Val Met Leu Pro Pro Gly Gln Val Val Gln Tyr Lys Tyr Val Asn Val
        355   360   365
Ala Ser Ser Gly Thr Pro Thr Trp Glu Lys Asp Pro Asn Arg Val Phe
    370   375   380
Thr Val Pro Ala Gly Cys Ala Thr Lys Gly Thr Val Gly Asp Thr Trp
385   390   395   400
Arg

<210> 467
<211> 2746
<212> DNA
<213> Cochliobolus heterostrophus ATCC 48331

<220>
<221> CDS
<222> (1)...(301)
<223> Exon

<221> CDS
<222> (380)...(454)
<223> Exon

<221> CDS
<222> (554)...(643)
<223> Exon

<221> CDS
<222> (717)...(2235)
<223> Exon

<221> CDS
<222> (2311)...(2746)
<223> Exon

<400> 467

```
atg atg aga cga gat gaa agc agt aac agt gac cgg gac gca tgc agt       48
Met Met Arg Arg Asp Glu Ser Ser Asn Ser Asp Arg Asp Ala Cys Ser
 1               5                  10                  15

tgc aag acg cta gcc tgt tcc agc ata gcc aac ctt ttt ggc ggc ttc       96
Cys Lys Thr Leu Ala Cys Ser Ser Ile Ala Asn Leu Phe Gly Gly Phe
            20                  25                  30

cct ccc tcc gga tgt cca agc agc gac gat ctt gcc cca aag tct tca      144
Pro Pro Ser Gly Cys Pro Ser Ser Asp Asp Leu Ala Pro Lys Ser Ser
         35                  40                  45

tcc aca aga ctt gca tcg agc tgg cag aca cgt cca ggg cac cat gtt      192
Ser Thr Arg Leu Ala Ser Ser Trp Gln Thr Arg Pro Gly His His Val
     50                  55                  60

ctt ttt tgc ttc tac acc tac ttg tct aca cca ctt cca cac cga cag      240
Leu Phe Cys Phe Tyr Thr Tyr Leu Ser Thr Pro Leu Pro His Arg Gln
 65                  70                  75                  80

ctg agt tgt acg att cat ggc agt ccg cgg cat ggc agt ttc gtg cag      288
Leu Ser Cys Thr Ile His Gly Ser Pro Arg His Gly Ser Phe Val Gln
                85                  90                  95

gct gac gtc act g gtgctgaatg ttgacaagca cggtctccac tcacatctag        341
Ala Asp Val Thr
            100

accattcaca acgcgcgggt ctcattcttg atagccag ca cag ccg aag gag tca     396
                                     Ala Gln Pro Lys Glu Ser
                                                    105

aca tac aaa cct gct cag cca tgg acc tat aca ggg tat tca ccc gct      444
Thr Tyr Lys Pro Ala Gln Pro Trp Thr Tyr Thr Gly Tyr Ser Pro Ala
            110                 115                 120

atg agc cgt a gtgggtataa aagtaaagaa catgagcgac aagtatggct            494
Met Ser Arg
        125

ctgtcgaatt ttccggtggg cttgaaaact tcgcttgaac attgatctcc ttggtccag     553
 ct gaa cca gta cgc tgc cta cca atc atc ggg aaa tcg tac aaa aca      600
Thr Glu Pro Val Arg Cys Leu Pro Ile Ile Gly Lys Ser Tyr Lys Thr
                130                 135                 140

gca cac ctc aag tct aaa ctg acc agc gca tcc ctg acg atg g            643
```

```
Ala His Leu Lys Ser Lys Leu Thr Ser Ala Ser Leu Thr Met
            145                 150                 155

gttcagaagg taggtcaagt ccagtgattg tttctcttcc gattggacta ctgcgatgct    703
caccactgaa cag aa cag aag cca acg cct gag aat agc acg ctg ttc cag    754
              Glu Gln Lys Pro Thr Pro Glu Asn Ser Thr Leu Phe Gln
                          160                 165

ggc ttc gag tgg aat gtc ccg gat gat cac aaa cac tgg aag cgt ctt      802
Gly Phe Glu Trp Asn Val Pro Asp Asp His Lys His Trp Lys Arg Leu
    170                 175                 180

gct gaa cag ctc cca aag ctc aaa gcg atc ggt atc act aac att tgg      850
Ala Glu Gln Leu Pro Lys Leu Lys Ala Ile Gly Ile Thr Asn Ile Trp
185                 190                 195                 200

cta cct ccg ggt tgt aaa gcc gcg aat cca aag ggt gtg ggc tat gac      898
Leu Pro Pro Gly Cys Lys Ala Ala Asn Pro Lys Gly Val Gly Tyr Asp
                205                 210                 215

atc tat gac cta tac gac ctg ggc gaa ttt gat cag aaa ggc gcc aaa      946
Ile Tyr Asp Leu Tyr Asp Leu Gly Glu Phe Asp Gln Lys Gly Ala Lys
            220                 225                 230

ggc acg aaa tgg gga aca aag gaa gag cta ttg gag ctc agc aag gtt      994
Gly Thr Lys Trp Gly Thr Lys Glu Glu Leu Leu Glu Leu Ser Lys Val
        235                 240                 245

gcc aag gaa cac aac gtt ggc ctt tac tgg gat gcc gtg ttg aat cac     1042
Ala Lys Glu His Asn Val Gly Leu Tyr Trp Asp Ala Val Leu Asn His
    250                 255                 260

aag gcg ggt gca gat agg aca gag aag tgc cgg gtt gtg gaa gtg gat     1090
Lys Ala Gly Ala Asp Arg Thr Glu Lys Cys Arg Val Val Glu Val Asp
265                 270                 275                 280

gaa aac gac cgc aca aag gag ata tcc gat gcc tac gaa atc gag ggg     1138
Glu Asn Asp Arg Thr Lys Glu Ile Ser Asp Ala Tyr Glu Ile Glu Gly
                285                 290                 295

tgg cta gga ttc gac ttc cca ggg cgc ggc gag cag tat tcc aag atg     1186
Trp Leu Gly Phe Asp Phe Pro Gly Arg Gly Glu Gln Tyr Ser Lys Met
            300                 305                 310

aag tac cac tgg gag cac ttt tct gga acc gac tac aac cag gcc aac     1234
Lys Tyr His Trp Glu His Phe Ser Gly Thr Asp Tyr Asn Gln Ala Asn
        315                 320                 325

gag aag aag gcc atc tac aag atc ctg ggt gaa aac aag ggc tgg tcg     1282
Glu Lys Lys Ala Ile Tyr Lys Ile Leu Gly Glu Asn Lys Gly Trp Ser
    330                 335                 340

caa agc gtg gac acg gaa tcc ggc aac gct gat tac atg atg ttc gcc     1330
Gln Ser Val Asp Thr Glu Ser Gly Asn Ala Asp Tyr Met Met Phe Ala
345                 350                 355                 360

gac att gac tac tcg cac cct gaa gtt caa gcg gat gtc aag aat tgg     1378
Asp Ile Asp Tyr Ser His Pro Glu Val Gln Ala Asp Val Lys Asn Trp
                365                 370                 375

ggt gtc tgg ata acc aaa gaa gtg gga ctc aag gga ttc cgc ctc gat     1426
Gly Val Trp Ile Thr Lys Glu Val Gly Leu Lys Gly Phe Arg Leu Asp
            380                 385                 390

gct gtg cag cac ttc agc tca aga ttt agc aac gaa tgg atc aag aac     1474
Ala Val Gln His Phe Ser Ser Arg Phe Ser Asn Glu Trp Ile Lys Asn
        395                 400                 405
```

```
ctc cgc gag cag tgc gga gac gac atg ttc atc gtg ggc gag ttc tgg     1522
Leu Arg Glu Gln Cys Gly Asp Asp Met Phe Ile Val Gly Glu Phe Trp
    410                 415                 420

tcg ggc gat gtc aag gaa atg agt gaa tgg ctc gac ggc atg cag cac     1570
Ser Gly Asp Val Lys Glu Met Ser Glu Trp Leu Asp Gly Met Gln His
425                 430                 435                 440

gaa att tgc ctg tac gac tcg ccc ttg ctt aac aag ttt ggc tcg ttg     1618
Glu Ile Cys Leu Tyr Asp Ser Pro Leu Leu Asn Lys Phe Gly Ser Leu
                445                 450                 455

tcg acg tgc gaa tcg gcc gat ctg agg aca gtg ttt gac gac tcg ctg     1666
Ser Thr Cys Glu Ser Ala Asp Leu Arg Thr Val Phe Asp Asp Ser Leu
            460                 465                 470

gtc cgg ctg cgg ccg atg gat gct gtg acg gtt gtc act aac cac gac     1714
Val Arg Leu Arg Pro Met Asp Ala Val Thr Val Val Thr Asn His Asp
        475                 480                 485

acg cag cct gga cag gtc atg gag acc aag att gag ggc ttc ttt aag     1762
Thr Gln Pro Gly Gln Val Met Glu Thr Lys Ile Glu Gly Phe Phe Lys
    490                 495                 500

ccg ctt gca tat gca ctg att ctg ctc cga cag gaa gga tac cca tgc     1810
Pro Leu Ala Tyr Ala Leu Ile Leu Leu Arg Gln Glu Gly Tyr Pro Cys
505                 510                 515                 520

ccg ttc tat ggc gat ctg tat ggc ctg tcg gag ccg cac gaa acg ccg     1858
Pro Phe Tyr Gly Asp Leu Tyr Gly Leu Ser Glu Pro His Glu Thr Pro
                525                 530                 535

gcg tcg tgc ggt ggc aag ctt gca gac ctt atc ctg gcg cgt aag ctg     1906
Ala Ser Cys Gly Gly Lys Leu Ala Asp Leu Ile Leu Ala Arg Lys Leu
            540                 545                 550

ttt gca tac gga gcg cag gaa gac tac ttg gac gag gcc aat tgc att     1954
Phe Ala Tyr Gly Ala Gln Glu Asp Tyr Leu Asp Glu Ala Asn Cys Ile
        555                 560                 565

ggg ttc gtg cgt cgt ggg aca tgg gac aag gtg gcg ggt ctt gca tgt     2002
Gly Phe Val Arg Arg Gly Thr Trp Asp Lys Val Ala Gly Leu Ala Cys
    570                 575                 580

gtg atg agc aat gcg ggg cca ggg cag atc aga atg gct gtg ggc gag     2050
Val Met Ser Asn Ala Gly Pro Gly Gln Ile Arg Met Ala Val Gly Glu
585                 590                 595                 600

atg cac gcc ggg gag aag tgg agc gac gtg ctt gga tgg gag cag ggc     2098
Met His Ala Gly Glu Lys Trp Ser Asp Val Leu Gly Trp Glu Gln Gly
                605                 610                 615

gag gta gag att gac agc gaa ggc tac gga gtg ttc agg tgt cct ggc     2146
Glu Val Glu Ile Asp Ser Glu Gly Tyr Gly Val Phe Arg Cys Pro Gly
            620                 625                 630

acc agc gtc gcg gtg tgg gtg cga agt gat gct gag gga cgc gac cag     2194
Thr Ser Val Ala Val Trp Val Arg Ser Asp Ala Glu Gly Arg Asp Gln
        635                 640                 645

ttc ccc acc aac tgg gat gcg gat gtg tat ggc cag gga tc gtgtgcgtaa   2245
Phe Pro Thr Asn Trp Asp Ala Asp Val Tyr Gly Gln Gly Ser
    650                 655                 660

tggttgcgat ggagggtgta cggatggggt agttgaatca cggcgctggc ggccgatgac   2305
gcaag g ccg ccc acc ctc ttc act tgt cca agt cag gcg tgt act gtt     2353
        Pro Pro Thr Leu Phe Thr Cys Pro Ser Gln Ala Cys Thr Val
                665                 670                 675
```

```
gtg agt agc gat ggg cat tat atc act gga aaa aga aga cat ggc gtg     2401
Val Ser Ser Asp Gly His Tyr Ile Thr Gly Lys Arg Arg His Gly Val
            680                 685                 690

atg cgc tgt cgt cgc cct ctg ccc tgt cgc cgc ggc tcg cag gct gtt     2449
Met Arg Cys Arg Arg Pro Leu Pro Cys Arg Arg Gly Ser Gln Ala Val
        695                 700                 705

agc tgg tgg aag atg cgc gaa agt gcg tgg ggc ggc cga cag ggc tgc     2497
Ser Trp Trp Lys Met Arg Glu Ser Ala Trp Gly Gly Arg Gln Gly Cys
    710                 715                 720

cgc gca gca cag cca atc acc cgc cgc aca aca ata gcc gtt ttg gac     2545
Arg Ala Ala Gln Pro Ile Thr Arg Arg Thr Thr Ile Ala Val Leu Asp
725                 730                 735                 740

gtt tcg cgg cct ttc tgg caa gcg agt atg cga cat ccc gcg aag cat     2593
Val Ser Arg Pro Phe Trp Gln Ala Ser Met Arg His Pro Ala Lys His
                745                 750                 755

gat cga cgg ccc atg ccg gca gaa tgc tcc ctc cct ctg tac cat gat     2641
Asp Arg Arg Pro Met Pro Ala Glu Cys Ser Leu Pro Leu Tyr His Asp
            760                 765                 770

gca cac gca cac gca cat gct cgc acg cgc gcc cgc gac acg gcc aac     2689
Ala His Ala His Ala His Ala Arg Thr Arg Ala Arg Asp Thr Ala Asn
        775                 780                 785

atg cga ggc caa cgg ccc cgt cct gcc gcc gga cat gaa gcc gca tgt     2737
Met Arg Gly Gln Arg Pro Arg Pro Ala Ala Gly His Glu Ala Ala Cys
    790                 795                 800

gcc atg tga                                                        2746
Ala Met  *
805
```

<210> 468
<211> 100
<212> PRT
<213> Cochliobolus heterostrophus ATCC 48331

<400> 468

```
Met Met Arg Arg Asp Glu Ser Ser Asn Ser Asp Arg Asp Ala Cys Ser
 1               5                  10                  15
Cys Lys Thr Leu Ala Cys Ser Ser Ile Ala Asn Leu Phe Gly Gly Phe
            20                  25                  30
Pro Pro Ser Gly Cys Pro Ser Ser Asp Asp Leu Ala Pro Lys Ser Ser
        35                  40                  45
Ser Thr Arg Leu Ala Ser Ser Trp Gln Thr Arg Pro Gly His His Val
    50                  55                  60
Leu Phe Cys Phe Tyr Thr Tyr Leu Ser Thr Pro Leu Pro His Arg Gln
65                  70                  75                  80
Leu Ser Cys Thr Ile His Gly Ser Pro Arg His Gly Ser Phe Val Gln
                85                  90                  95
Ala Asp Val Thr
            100
```

<210> 469
<211> 25
<212> PRT
<213> Cochliobolus heterostrophus ATCC 48331

<400> 469

```
Ala Gln Pro Lys Glu Ser Thr Tyr Lys Pro Ala Gln Pro Trp Thr Tyr
 1               5                  10                  15
Thr Gly Tyr Ser Pro Ala Met Ser Arg
```

20 25

<210> 470
<211> 30
<212> PRT
<213> Cochliobolus heterostrophus ATCC 48331

<400> 470
Thr Glu Pro Val Arg Cys Leu Pro Ile Ile Gly Lys Ser Tyr Lys Thr
1 5 10 15
Ala His Leu Lys Ser Lys Leu Thr Ser Ala Ser Leu Thr Met
20 25 30

<210> 471
<211> 507
<212> PRT
<213> Cochliobolus heterostrophus ATCC 48331

<400> 471
Glu Gln Lys Pro Thr Pro Glu Asn Ser Thr Leu Phe Gln Gly Phe Glu
1 5 10 15
Trp Asn Val Pro Asp Asp His Lys His Trp Lys Arg Leu Ala Glu Gln
20 25 30
Leu Pro Lys Leu Lys Ala Ile Gly Ile Thr Asn Ile Trp Leu Pro Pro
35 40 45
Gly Cys Lys Ala Ala Asn Pro Lys Gly Val Gly Tyr Asp Ile Tyr Asp
50 55 60
Leu Tyr Asp Leu Gly Glu Phe Asp Gln Lys Gly Ala Lys Gly Thr Lys
65 70 75 80
Trp Gly Thr Lys Glu Glu Leu Leu Glu Leu Ser Lys Val Ala Lys Glu
85 90 95
His Asn Val Gly Leu Tyr Trp Asp Ala Val Leu Asn His Lys Ala Gly
100 105 110
Ala Asp Arg Thr Glu Lys Cys Arg Val Val Glu Val Asp Glu Asn Asp
115 120 125
Arg Thr Lys Glu Ile Ser Asp Ala Tyr Glu Ile Glu Gly Trp Leu Gly
130 135 140
Phe Asp Phe Pro Gly Arg Gly Glu Gln Tyr Ser Lys Met Lys Tyr His
145 150 155 160
Trp Glu His Phe Ser Gly Thr Asp Tyr Asn Gln Ala Asn Glu Lys Lys
165 170 175
Ala Ile Tyr Lys Ile Leu Gly Glu Asn Lys Gly Trp Ser Gln Ser Val
180 185 190
Asp Thr Glu Ser Gly Asn Ala Asp Tyr Met Met Phe Ala Asp Ile Asp
195 200 205
Tyr Ser His Pro Glu Val Gln Ala Asp Val Lys Asn Trp Gly Val Trp
210 215 220
Ile Thr Lys Glu Val Gly Leu Lys Gly Phe Arg Leu Asp Ala Val Gln
225 230 235 240
His Phe Ser Ser Arg Phe Ser Asn Glu Trp Ile Lys Asn Leu Arg Glu
245 250 255
Gln Cys Gly Asp Asp Met Phe Ile Val Gly Glu Phe Trp Ser Gly Asp
260 265 270
Val Lys Glu Met Ser Glu Trp Leu Asp Gly Met Gln His Glu Ile Cys
275 280 285
Leu Tyr Asp Ser Pro Leu Leu Asn Lys Phe Gly Ser Leu Ser Thr Cys
290 295 300
Glu Ser Ala Asp Leu Arg Thr Val Phe Asp Asp Ser Leu Val Arg Leu
305 310 315 320
Arg Pro Met Asp Ala Val Thr Val Val Thr Asn His Asp Thr Gln Pro
325 330 335
Gly Gln Val Met Glu Thr Lys Ile Glu Gly Phe Phe Lys Pro Leu Ala
340 345 350
Tyr Ala Leu Ile Leu Leu Arg Gln Glu Gly Tyr Pro Cys Pro Phe Tyr
355 360 365
Gly Asp Leu Tyr Gly Leu Ser Glu Pro His Glu Thr Pro Ala Ser Cys
370 375 380
Gly Gly Lys Leu Ala Asp Leu Ile Leu Ala Arg Lys Leu Phe Ala Tyr

385 390 395 400
Gly Ala Gln Glu Asp Tyr Leu Asp Glu Ala Asn Cys Ile Gly Phe Val
405 410 415
Arg Arg Gly Thr Trp Asp Lys Val Ala Gly Leu Ala Cys Val Met Ser
420 425 430
Asn Ala Gly Pro Gly Gln Ile Arg Met Ala Val Gly Glu Met His Ala
435 440 445
Gly Glu Lys Trp Ser Asp Val Leu Gly Trp Glu Gln Gly Glu Val Glu
450 455 460
Ile Asp Ser Glu Gly Tyr Gly Val Phe Arg Cys Pro Gly Thr Ser Val
465 470 475 480
Ala Val Trp Val Arg Ser Asp Ala Glu Gly Arg Asp Gln Phe Pro Thr
485 490 495
Asn Trp Asp Ala Asp Val Tyr Gly Gln Gly Ser
500 505

<210> 472
<211> 144
<212> PRT
<213> Cochliobolus heterostrophus ATCC 48331

<400> 472
Pro Pro Thr Leu Phe Thr Cys Pro Ser Gln Ala Cys Thr Val Val Ser
1 5 10 15
Ser Asp Gly His Tyr Ile Thr Gly Lys Arg Arg His Gly Val Met Arg
20 25 30
Cys Arg Arg Pro Leu Pro Cys Arg Arg Gly Ser Gln Ala Val Ser Trp
35 40 45
Trp Lys Met Arg Glu Ser Ala Trp Gly Gly Arg Gln Gly Cys Arg Ala
50 55 60
Ala Gln Pro Ile Thr Arg Arg Thr Thr Ile Ala Val Leu Asp Val Ser
65 70 75 80
Arg Pro Phe Trp Gln Ala Ser Met Arg His Pro Ala Lys His Asp Arg
85 90 95
Arg Pro Met Pro Ala Glu Cys Ser Leu Pro Leu Tyr His Asp Ala His
100 105 110
Ala His Ala His Ala Arg Thr Arg Ala Arg Asp Thr Ala Asn Met Arg
115 120 125
Gly Gln Arg Pro Arg Pro Ala Ala Gly His Glu Ala Ala Cys Ala Met
130 135 140

<210> 473
<211> 2421
<212> DNA
<213> Cochliobolus heterostrophus ATCC 48331

<400> 473
atgatgagac gagatgaaag cagtaacagt gaccgggacg catgcagttg caagacgcta 60
gcctgttcca gcatagccaa ccttttggc ggcttccctc cctccggatg tccaagcagc 120
gacgatcttg ccccaaagtc ttcatccaca agacttgcat cgagctggca gacacgtcca 180
gggcaccatg ttctttttg cttctacacc tacttgtcta caccacttcc acaccgacag 240
ctgagttgta cgattcatgg cagtccgcgg catggcagtt tcgtgcaggc tgacgtcact 300
gcacagccga aggagtcaac atacaaacct gctcagccat ggacctatac agggtattca 360
cccgctatga gccgtactga accagtacgc tgcctaccaa tcatcgggaa atcgtacaaa 420
acagcacacc tcaagtctaa actgaccagc gcatccctga cgatggaaca gaagccaacg 480
cctgagaata gcacgctgtt ccagggcttc gagtggaatg tcccggatga tcacaaacac 540
tggaagcgtc ttgctgaaca gctcccaaag ctcaaagcga tcggtatcac taacatttgg 600
ctacctccgg gttgtaaagc cgcgaatcca aagggtgtgg gctatgacat ctatgaccta 660
tacgacctgg gcgaatttga tcagaaaggc gccaaaggca cgaaatgggg aacaaaggaa 720
gagctattgg agctcagcaa ggttgccaag gaacacaacg ttggccttta ctgggatgcc 780
gtgttgaatc acaaggcggg tgcagatagg acagagaagt gccgggttgt ggaagtggat 840
gaaaacgacc gcacaaagga gatatccgat gcctacgaaa tcgaggggtg gctaggattc 900
gacttcccag ggcgcggcga gcagtattcc aagatgaagt accactggga gcacttttct 960
ggaaccgact acaaccaggc caacgagaag aaggccatct acaagatcct gggtgaaaac 1020
aagggctggt cgcaaagcgt ggacacggaa tccggcaacg ctgattacat gatgttcgcc 1080
gacattgact actcgcaccc tgaagttcaa gcggatgtca agaattgggg tgtctggata 1140
accaaagaag tgggactcaa gggattccgc ctcgatgctg tgcagcactt cagctcaaga 1200
tttagcaacg aatggatcaa gaacctccgc gagcagtgcg gagacgacat gttcatcgtg 1260

ggcgagttct ggtcgggcga tgtcaaggaa atgagtgaat ggctcgacgg catgcagcac 1320
gaaatttgcc tgtacgactc gcccttgctt aacaagtttg gctcgttgtc gacgtgcgaa 1380
tcggccgatc tgaggacagt gtttgacgac tcgctggtcc ggctgcggcc gatggatgct 1440
gtgacggttg tcactaacca cgacacgcag cctggacagg tcatggagac caagattgag 1500
ggcttcttta agccgcttgc atatgcactg attctgctcc gacaggaagg atacccatgc 1560
ccgttctatg gcgatctgta tggcctgtcg gagccgcacg aaacgccggc gtcgtgcggt 1620
ggcaagcttg cagaccttat cctggcgcgt aagctgtttg catacggagc gcaggaagac 1680
tacttggacg aggccaattg cattgggttc gtgcgtcgtg ggacatggga caaggtggcg 1740
ggtcttgcat gtgtgatgag caatgcgggg ccagggcaga tcagaatggc tgtgggcgag 1800
atgcacgccg gggagaagtg gagcgacgtg cttggatggg agcagggcga ggtagagatt 1860
gacagcgaag gctacggagt gttcaggtgt cctggcacca gcgtcgcggt gtgggtgcga 1920
agtgatgctg agggacgcga ccagttcccc accaactggg atgcggatgt gtatggccag 1980
ggatcgccgc ccaccctctt cacttgtcca agtcaggcgt gtactgttgt gagtagcgat 2040
gggcattata tcactggaaa aagaagacat ggcgtgatgc gctgtcgtcg ccctctgccc 2100
tgtcgccgcg gctcgcaggc tgttagctgg tggaagatgc gcgaaagtgc gtggggcggc 2160
cgacagggct gccgcgcagc acagccaatc acccgccgca caacaatagc cgttttggac 2220
gtttcgcggc ctttctggca agcgagtatg cgacatcccg cgaagcatga tcgacggccc 2280
atgccggcag aatgctccct ccctctgtac catgatgcac acgcacacgc acatgctcgc 2340
acgcgcgccc gcgacacggc caacatgcga ggccaacggc cccgtcctgc cgccggacat 2400
gaagccgcat gtgccatgtg a 2421

<210> 474
<211> 806
<212> PRT
<213> Cochliobolus heterostrophus ATCC 48331

<220>
<221> DOMAIN
<222> (164)...(550)
<223> Alpha amylase, catalytic domain

<400> 474
Met Met Arg Arg Asp Glu Ser Ser Asn Ser Asp Arg Asp Ala Cys Ser
 1   5   10   15
Cys Lys Thr Leu Ala Cys Ser Ser Ile Ala Asn Leu Phe Gly Gly Phe
            20   25   30
Pro Pro Ser Gly Cys Pro Ser Ser Asp Asp Leu Ala Pro Lys Ser Ser
        35   40   45
Ser Thr Arg Leu Ala Ser Ser Trp Gln Thr Arg Pro Gly His His Val
    50   55   60
Leu Phe Cys Phe Tyr Thr Tyr Leu Ser Thr Pro Leu Pro His Arg Gln
65   70   75   80
Leu Ser Cys Thr Ile His Gly Ser Pro Arg His Gly Ser Phe Val Gln
                85   90   95
Ala Asp Val Thr Ala Gln Pro Lys Glu Ser Thr Tyr Lys Pro Ala Gln
            100   105   110
Pro Trp Thr Tyr Thr Gly Tyr Ser Pro Ala Met Ser Arg Thr Glu Pro
        115   120   125
Val Arg Cys Leu Pro Ile Ile Gly Lys Ser Tyr Lys Thr Ala His Leu
    130   135   140
Lys Ser Lys Leu Thr Ser Ala Ser Leu Thr Met Glu Gln Lys Pro Thr
145   150   155   160
Pro Glu Asn Ser Thr Leu Phe Gln Gly Phe Glu Trp Asn Val Pro Asp
                165   170   175
Asp His Lys His Trp Lys Arg Leu Ala Glu Gln Leu Pro Lys Leu Lys
            180   185   190
Ala Ile Gly Ile Thr Asn Ile Trp Leu Pro Pro Gly Cys Lys Ala Ala
        195   200   205
Asn Pro Lys Gly Val Gly Tyr Asp Ile Tyr Asp Leu Tyr Asp Leu Gly
    210   215   220
Glu Phe Asp Gln Lys Gly Ala Lys Gly Thr Lys Trp Gly Thr Lys Glu
225   230   235   240
Glu Leu Leu Glu Leu Ser Lys Val Ala Lys Glu His Asn Val Gly Leu
                245   250   255
Tyr Trp Asp Ala Val Leu Asn His Lys Ala Gly Ala Asp Arg Thr Glu
            260   265   270
Lys Cys Arg Val Val Glu Val Asp Glu Asn Asp Arg Thr Lys Glu Ile
        275   280   285

```
Ser Asp Ala Tyr Glu Ile Glu Gly Trp Leu Gly Phe Asp Phe Pro Gly
    290                 295                 300
Arg Gly Glu Gln Tyr Ser Lys Met Lys Tyr His Trp Glu His Phe Ser
305                 310                 315                 320
Gly Thr Asp Tyr Asn Gln Ala Asn Glu Lys Lys Ala Ile Tyr Lys Ile
                325                 330                 335
Leu Gly Glu Asn Lys Gly Trp Ser Gln Ser Val Asp Thr Glu Ser Gly
            340                 345                 350
Asn Ala Asp Tyr Met Met Phe Ala Asp Ile Asp Tyr Ser His Pro Glu
        355                 360                 365
Val Gln Ala Asp Val Lys Asn Trp Gly Val Trp Ile Thr Lys Glu Val
    370                 375                 380
Gly Leu Lys Gly Phe Arg Leu Asp Ala Val Gln His Phe Ser Ser Arg
385                 390                 395                 400
Phe Ser Asn Glu Trp Ile Lys Asn Leu Arg Glu Gln Cys Gly Asp Asp
                405                 410                 415
Met Phe Ile Val Gly Glu Phe Trp Ser Gly Asp Val Lys Glu Met Ser
            420                 425                 430
Glu Trp Leu Asp Gly Met Gln His Glu Ile Cys Leu Tyr Asp Ser Pro
        435                 440                 445
Leu Leu Asn Lys Phe Gly Ser Leu Ser Thr Cys Glu Ser Ala Asp Leu
    450                 455                 460
Arg Thr Val Phe Asp Asp Ser Leu Val Arg Leu Arg Pro Met Asp Ala
465                 470                 475                 480
Val Thr Val Val Thr Asn His Asp Thr Gln Pro Gly Gln Val Met Glu
                485                 490                 495
Thr Lys Ile Glu Gly Phe Phe Lys Pro Leu Ala Tyr Ala Leu Ile Leu
            500                 505                 510
Leu Arg Gln Glu Gly Tyr Pro Cys Pro Phe Tyr Gly Asp Leu Tyr Gly
        515                 520                 525
Leu Ser Glu Pro His Glu Thr Pro Ala Ser Cys Gly Gly Lys Leu Ala
    530                 535                 540
Asp Leu Ile Leu Ala Arg Lys Leu Phe Ala Tyr Gly Ala Gln Glu Asp
545                 550                 555                 560
Tyr Leu Asp Glu Ala Asn Cys Ile Gly Phe Val Arg Arg Gly Thr Trp
                565                 570                 575
Asp Lys Val Ala Gly Leu Ala Cys Val Met Ser Asn Ala Gly Pro Gly
            580                 585                 590
Gln Ile Arg Met Ala Val Gly Glu Met His Ala Gly Glu Lys Trp Ser
        595                 600                 605
Asp Val Leu Gly Trp Glu Gln Gly Glu Val Glu Ile Asp Ser Glu Gly
    610                 615                 620
Tyr Gly Val Phe Arg Cys Pro Gly Thr Ser Val Ala Val Trp Val Arg
625                 630                 635                 640
Ser Asp Ala Glu Gly Arg Asp Gln Phe Pro Thr Asn Trp Asp Ala Asp
                645                 650                 655
Val Tyr Gly Gln Gly Ser Pro Pro Thr Leu Phe Thr Cys Pro Ser Gln
            660                 665                 670
Ala Cys Thr Val Val Ser Ser Asp Gly His Tyr Ile Thr Gly Lys Arg
        675                 680                 685
Arg His Gly Val Met Arg Cys Arg Arg Pro Leu Pro Cys Arg Arg Gly
    690                 695                 700
Ser Gln Ala Val Ser Trp Trp Lys Met Arg Glu Ser Ala Trp Gly Gly
705                 710                 715                 720
Arg Gln Gly Cys Arg Ala Ala Gln Pro Ile Thr Arg Arg Thr Thr Ile
                725                 730                 735
Ala Val Leu Asp Val Ser Arg Pro Phe Trp Gln Ala Ser Met Arg His
            740                 745                 750
Pro Ala Lys His Asp Arg Arg Pro Met Pro Ala Glu Cys Ser Leu Pro
        755                 760                 765
Leu Tyr His Asp Ala His Ala His Ala His Ala Arg Thr Arg Ala Arg
    770                 775                 780
Asp Thr Ala Asn Met Arg Gly Gln Arg Pro Arg Pro Ala Ala Gly His
785                 790                 795                 800
Glu Ala Ala Cys Ala Met
                805
```

<210> 475

<211> 7181
<212> DNA
<213> Fungi

<220>
<221> CDS
<222> (1)...(2734)
<223> Exon

<221> CDS
<222> (2785)...(7181)
<223> Exon

<400> 475

```
atg caa acc caa tta cga cat gga gga gat cta cag ggt ctc att gat       48
Met Gln Thr Gln Leu Arg His Gly Gly Asp Leu Gln Gly Leu Ile Asp
 1               5                  10                  15

agt ctg gac tat ata caa gga atg gga ata aag gga att tat att gct       96
Ser Leu Asp Tyr Ile Gln Gly Met Gly Ile Lys Gly Ile Tyr Ile Ala
            20                  25                  30

gga agt cct atg atc aat gcg cct tgg ggc gct gat caa tac tct cct      144
Gly Ser Pro Met Ile Asn Ala Pro Trp Gly Ala Asp Gln Tyr Ser Pro
        35                  40                  45

ctg gat ttc agt ctt ctc gac caa cat ttc gga gat atc gat ttg tgg      192
Leu Asp Phe Ser Leu Leu Asp Gln His Phe Gly Asp Ile Asp Leu Trp
    50                  55                  60

agg aaa act acc act gcc atc cat gac agg gga atg tat gtg atg ctc      240
Arg Lys Thr Thr Thr Ala Ile His Asp Arg Gly Met Tyr Val Met Leu
65                  70                  75                  80

gac aac acc atg gga acc atg gga gac ttg att gct ttc aaa ggc tat      288
Asp Asn Thr Met Gly Thr Met Gly Asp Leu Ile Ala Phe Lys Gly Tyr
                85                  90                  95

gaa aat tct tcg act ccg ttc acc ctg acc gag cac gaa gtc ctt tgg      336
Glu Asn Ser Ser Thr Pro Phe Thr Leu Thr Glu His Glu Val Leu Trp
            100                 105                 110

cgt aat gac aac cgc caa tat ctg gat ttt tcc ttc ggg tcg aag tac      384
Arg Asn Asp Asn Arg Gln Tyr Leu Asp Phe Ser Phe Gly Ser Lys Tyr
        115                 120                 125

aat gag acc tgc aac tac cct aga ttc tgg ctt gat act gga tat acc      432
Asn Glu Thr Cys Asn Tyr Pro Arg Phe Trp Leu Asp Thr Gly Tyr Thr
    130                 135                 140

gtc aag gaa gat gtc act gac ctc atg aat ggg tgc tat gat tcc gaa      480
Val Lys Glu Asp Val Thr Asp Leu Met Asn Gly Cys Tyr Asp Ser Glu
145                 150                 155                 160

ttc gat caa tat ggt gat acc gaa gct ttt ggt gtc ttc cct gat tgg      528
Phe Asp Gln Tyr Gly Asp Thr Glu Ala Phe Gly Val Phe Pro Asp Trp
                165                 170                 175

caa cgt cag ttg tca aag ttc gcc tct gtc cag gat cgt ttg aga gaa      576
Gln Arg Gln Leu Ser Lys Phe Ala Ser Val Gln Asp Arg Leu Arg Glu
            180                 185                 190

tgg gtt cca tcc gta aga act aaa atc gag atc ttc act tgc atg atc      624
Trp Val Pro Ser Val Arg Thr Lys Ile Glu Ile Phe Thr Cys Met Ile
        195                 200                 205

atc caa caa ttg gac att gat gga ctg cgt gtg gac aag gcg aca caa      672
Ile Gln Gln Leu Asp Ile Asp Gly Leu Arg Val Asp Lys Ala Thr Gln
```

```
    210                 215                 220

att acg gtc gat gca ttg gga agt ttc agt cac gca atc aga gag tgc      720
Ile Thr Val Asp Ala Leu Gly Ser Phe Ser His Ala Ile Arg Glu Cys
225                 230                 235                 240

gca aga gat gtg aac aag aca aat ttc atg gtt act ggt gaa atc act      768
Ala Arg Asp Val Asn Lys Thr Asn Phe Met Val Thr Gly Glu Ile Thr
                245                 250                 255

ggt gga aac act ttt gga agt att tac ctc ggg cgt gga cgt caa cct      816
Gly Gly Asn Thr Phe Gly Ser Ile Tyr Leu Gly Arg Gly Arg Gln Pro
            260                 265                 270

gac atg tta ccg tcc agc atg gca gat gcg gtt aaa ttg acc tcc aat      864
Asp Met Leu Pro Ser Ser Met Ala Asp Ala Val Lys Leu Thr Ser Asn
        275                 280                 285

tct agt ctg gca tat ttt atc cgt gat cac gat caa aac gcg ctc gac      912
Ser Ser Leu Ala Tyr Phe Ile Arg Asp His Asp Gln Asn Ala Leu Asp
    290                 295                 300

tct gct gct ttt cac tat agt gtt tat cgc aca ctg acc aga ttt ctg      960
Ser Ala Ala Phe His Tyr Ser Val Tyr Arg Thr Leu Thr Arg Phe Leu
305                 310                 315                 320

gga atg gac gga aac ttg gca agt ggt tat gat aca cca act aac tgg     1008
Gly Met Asp Gly Asn Leu Ala Ser Gly Tyr Asp Thr Pro Thr Asn Trp
                325                 330                 335

gta gat gcg tgg aac gaa atg ctt gta acg aac gat ttt gtt aac cct     1056
Val Asp Ala Trp Asn Glu Met Leu Val Thr Asn Asp Phe Val Asn Pro
            340                 345                 350

aat act gga gtg ttc gat cca aga cat atg tac gga gtc agc aat cag     1104
Asn Thr Gly Val Phe Asp Pro Arg His Met Tyr Gly Val Ser Asn Gln
        355                 360                 365

gat gtt ttc aga tgg cca gct atc cag aat ggc aca caa aaa caa ctt     1152
Asp Val Phe Arg Trp Pro Ala Ile Gln Asn Gly Thr Gln Lys Gln Leu
    370                 375                 380

cta ggg ttc ttc atc acg act ata cac atg cct ggt att cct ttg ctt     1200
Leu Gly Phe Phe Ile Thr Thr Ile His Met Pro Gly Ile Pro Leu Leu
385                 390                 395                 400

ttg tgg ggt gaa gaa caa gca ttc tat gtc tta gac aat aca gcc gac     1248
Leu Trp Gly Glu Glu Gln Ala Phe Tyr Val Leu Asp Asn Thr Ala Asp
                405                 410                 415

aat tat ctc ttc ggt cga caa gct atg tct cct tct caa gca tgg cag     1296
Asn Tyr Leu Phe Gly Arg Gln Ala Met Ser Pro Ser Gln Ala Trp Gln
            420                 425                 430

act cac gga tgc tac gcc ctt ggc agt ggt aca tat tat caa ttc ccc     1344
Thr His Gly Cys Tyr Ala Leu Gly Ser Gly Thr Tyr Tyr Gln Phe Pro
        435                 440                 445

aat gac tct tgc act cga ggt tgc aat gat gat tgg aac agc cgc gat     1392
Asn Asp Ser Cys Thr Arg Gly Cys Asn Asp Asp Trp Asn Ser Arg Asp
    450                 455                 460

cac aga gat cca tct cat cca gtt aga aat att gtc aag gca atg tac     1440
His Arg Asp Pro Ser His Pro Val Arg Asn Ile Val Lys Ala Met Tyr
465                 470                 475                 480

cag atg cga gac aac tat cca gtt ttg acc gat gga gca ttt tta cga     1488
Gln Met Arg Asp Asn Tyr Pro Val Leu Thr Asp Gly Ala Phe Leu Arg
```

```
                485                 490                 495

caa ctt tcc aac cag aca atc gaa atc ctc tta cca ggt agt aac gga     1536
Gln Leu Ser Asn Gln Thr Ile Glu Ile Leu Leu Pro Gly Ser Asn Gly
            500                 505                 510

aca gct act gaa act gga atg tgg agt acc atg cga ggt caa tac gat     1584
Thr Ala Thr Glu Thr Gly Met Trp Ser Thr Met Arg Gly Gln Tyr Asp
        515                 520                 525

ggt att caa gat tta tct aaa gct ggt ggt caa gca aac cag tct gtt     1632
Gly Ile Gln Asp Leu Ser Lys Ala Gly Gly Gln Ala Asn Gln Ser Val
    530                 535                 540

tgg atg gtt tac cac aac aat tat gac aca gtt aca tat tct ttt gat     1680
Trp Met Val Tyr His Asn Asn Tyr Asp Thr Val Thr Tyr Ser Phe Asp
545                 550                 555                 560

tgc agt agc aac gaa acc tcg ctt gtt gct cct ttc ctt gac ggt gtc     1728
Cys Ser Ser Asn Glu Thr Ser Leu Val Ala Pro Phe Leu Asp Gly Val
                565                 570                 575

act gtc aaa aat ttg ttc tat ccg tat gag gag gtt aca ttg aaa gca     1776
Thr Val Lys Asn Leu Phe Tyr Pro Tyr Glu Glu Val Thr Leu Lys Ala
            580                 585                 590

ggt cct ttt cag cta gga att gca aat cag act ggc ttc aat ggc tgt     1824
Gly Pro Phe Gln Leu Gly Ile Ala Asn Gln Thr Gly Phe Asn Gly Cys
        595                 600                 605

ctc gat agt ctc act ctc gtt cca tac gag tat aaa gcg tat gtg cca     1872
Leu Asp Ser Leu Thr Leu Val Pro Tyr Glu Tyr Lys Ala Tyr Val Pro
    610                 615                 620

aaa aag agc tgg gtt gga cct tcg cct atg att acc tca ttt tca cca     1920
Lys Lys Ser Trp Val Gly Pro Ser Pro Met Ile Thr Ser Phe Ser Pro
625                 630                 635                 640

gga cac gat gca cga ctt gtt tct aca gtt gct ccc gga gaa caa gaa     1968
Gly His Asp Ala Arg Leu Val Ser Thr Val Ala Pro Gly Glu Gln Glu
                645                 650                 655

acc att gaa att gaa ttt ggc ttc agt gcc gaa atg aag tgc tct gat     2016
Thr Ile Glu Ile Glu Phe Gly Phe Ser Ala Glu Met Lys Cys Ser Asp
            660                 665                 670

ctc att tcg aaa atg gta att acc tcg gca aca gaa gac gga cag gtt     2064
Leu Ile Ser Lys Met Val Ile Thr Ser Ala Thr Glu Asp Gly Gln Val
        675                 680                 685

gca aca cta gcc aat gcc acc ggt gaa tat tct tgt act ttg ctt cca     2112
Ala Thr Leu Ala Asn Ala Thr Gly Glu Tyr Ser Cys Thr Leu Leu Pro
    690                 695                 700

agt gcg gtc aat gta acc gaa ttc gtc gga tca att cca acc atg tgg     2160
Ser Ala Val Asn Val Thr Glu Phe Val Gly Ser Ile Pro Thr Met Trp
705                 710                 715                 720

act ttc aaa gga aat ctt acc aat gta tcc aac ggt gtc cac tca atc     2208
Thr Phe Lys Gly Asn Leu Thr Asn Val Ser Asn Gly Val His Ser Ile
                725                 730                 735

act cta gcc aat atc acc agc aac tct gga aat tcc tca act gga gcc     2256
Thr Leu Ala Asn Ile Thr Ser Asn Ser Gly Asn Ser Ser Thr Gly Ala
            740                 745                 750

aca gat aca ttc ttg ttc cgt gtt gga caa gtt gat aat cct atc gct     2304
Thr Asp Thr Phe Leu Phe Arg Val Gly Gln Val Asp Asn Pro Ile Ala
```

```
        755                 760                 765

ttc ccc ggt tcc gca aac ttc aca aga gaa tta ctt cac aag tat gac     2352
Phe Pro Gly Ser Ala Asn Phe Thr Arg Glu Leu Leu His Lys Tyr Asp
    770                 775                 780

aat ggc aca ttg tat gtc agc cac aaa gct gct gga gca gat cgg ttc     2400
Asn Gly Thr Leu Tyr Val Ser His Lys Ala Ala Gly Ala Asp Arg Phe
785                 790                 795                 800

aaa tat tct ttg aac tgg gga agt tca tgg tct gaa tgg tta cca tat     2448
Lys Tyr Ser Leu Asn Trp Gly Ser Ser Trp Ser Glu Trp Leu Pro Tyr
                805                 810                 815

gtt ggt gga aat acc aca ctt gaa cca cag gct tgg act ggt aca agc     2496
Val Gly Gly Asn Thr Thr Leu Glu Pro Gln Ala Trp Thr Gly Thr Ser
            820                 825                 830

aaa caa aaa tgg gac gga gat cac gtc atc atg caa tat tgg agc aaa     2544
Lys Gln Lys Trp Asp Gly Asp His Val Ile Met Gln Tyr Trp Ser Lys
        835                 840                 845

ctt act ggt agc agt gat gta gtt caa cat gct gat cta gaa acc aac     2592
Leu Thr Gly Ser Ser Asp Val Val Gln His Ala Asp Leu Glu Thr Asn
    850                 855                 860

aaa tgg tcg cgt cgc ctt ccg cac ttg ttt gct aat ggt cca tac aac     2640
Lys Trp Ser Arg Arg Leu Pro His Leu Phe Ala Asn Gly Pro Tyr Asn
865                 870                 875                 880

cag tat gga tat gat act ggt atc gat aat ggt ttc act ttg aac aag     2688
Gln Tyr Gly Tyr Asp Thr Gly Ile Asp Asn Gly Phe Thr Leu Asn Lys
                885                 890                 895

aaa gac agt ctc tgg gaa tat aag ttt atg act gaa tgg cca gcc g       2734
Lys Asp Ser Leu Trp Glu Tyr Lys Phe Met Thr Glu Trp Pro Ala
            900                 905                 910

gtaagttgta tcttcaatcc gatgtatcaa ggcactctaa caatttctag tc atg caa   2792
                                                        Val Met Gln


gtg aat gtt tgg ggt atg aat cct gat ggg caa cct gac gaa act atc     2840
Val Asn Val Trp Gly Met Asn Pro Asp Gly Gln Pro Asp Glu Thr Ile
915                 920                 925                 930

gtg ttg gga gat atc gac aat gat aat atc ctt gat cgt atg cca cct     2888
Val Leu Gly Asp Ile Asp Asn Asp Asn Ile Leu Asp Arg Met Pro Pro
                935                 940                 945

tca tcg cta tcc acg gcg ctc atc aac atc aca cgc gtg cca cca tca     2936
Ser Ser Leu Ser Thr Ala Leu Ile Asn Ile Thr Arg Val Pro Pro Ser
            950                 955                 960

cca tat ctc gga tac aaa att gca atc gac gac ggt aat tat aga tac     2984
Pro Tyr Leu Gly Tyr Lys Ile Ala Ile Asp Asp Gly Asn Tyr Arg Tyr
        965                 970                 975

tat caa atc cct tac gga aat cgt tcc tac caa atg ttg atg gtt gtc     3032
Tyr Gln Ile Pro Tyr Gly Asn Arg Ser Tyr Gln Met Leu Met Val Val
    980                 985                 990

ctt ttc tgg att gtc cct gta ctc agt ggt ggt ctc agt atc tgg ctg     3080
Leu Phe Trp Ile Val Pro Val Leu Ser Gly Gly Leu Ser Ile Trp Leu
 995                1000                1005                1010

ttc atg aaa tct ttc tac gcc gtc aaa ctc aac agt gtc ggt gtc aag     3128
Phe Met Lys Ser Phe Tyr Ala Val Lys Leu Asn Ser Val Gly Val Lys
```

```
                1015                1020                1025

gag gca aag aac tat atc cca tta gcc ttg cgc aga aag ctg aag aga     3176
Glu Ala Lys Asn Tyr Ile Pro Leu Ala Leu Arg Arg Lys Leu Lys Arg
            1030                1035                1040

gaa aag aaa gag gac cca tct gag aaa gaa atg acc gtc atg gga gga     3224
Glu Lys Lys Glu Asp Pro Ser Glu Lys Glu Met Thr Val Met Gly Gly
        1045                1050                1055

ttg cac tca atg ttc cac gaa aat gcc aca caa ctt gga ctg cag gct     3272
Leu His Ser Met Phe His Glu Asn Ala Thr Gln Leu Gly Leu Gln Ala
    1060                1065                1070

gat gtt gca gcg acc gga cga cga act gta ctt att gct act atg gaa     3320
Asp Val Ala Ala Thr Gly Arg Arg Thr Val Leu Ile Ala Thr Met Glu
1075                1080                1085                1090

tat gat att gaa gat tgg gct atc aag atc aag att ggt ggt ctt ggt     3368
Tyr Asp Ile Glu Asp Trp Ala Ile Lys Ile Lys Ile Gly Gly Leu Gly
                1095                1100                1105

gtc atg gca caa ttg atg gga aag aat ttg ggt cat caa gat ctc att     3416
Val Met Ala Gln Leu Met Gly Lys Asn Leu Gly His Gln Asp Leu Ile
            1110                1115                1120

tgg gtc gtt cct tgt gtt ggt ggt atc gat tat cct atc gat cac caa     3464
Trp Val Val Pro Cys Val Gly Gly Ile Asp Tyr Pro Ile Asp His Gln
        1125                1130                1135

gct gaa cct atg gat gtc aca att ctt gga aat tcc tac act gtt caa     3512
Ala Glu Pro Met Asp Val Thr Ile Leu Gly Asn Ser Tyr Thr Val Gln
    1140                1145                1150

gtt cag tac cat act cta cga aac atc act tac gtt ctt ctc gat gca     3560
Val Gln Tyr His Thr Leu Arg Asn Ile Thr Tyr Val Leu Leu Asp Ala
1155                1160                1165                1170

cca gtt ttc aga gca caa tct aag tct gag cca tac cca gct cgt atg     3608
Pro Val Phe Arg Ala Gln Ser Lys Ser Glu Pro Tyr Pro Ala Arg Met
                1175                1180                1185

gat gat tta gac tcg gcc gtt tac tac agt gct tgg aat caa tgt atc     3656
Asp Asp Leu Asp Ser Ala Val Tyr Tyr Ser Ala Trp Asn Gln Cys Ile
            1190                1195                1200

gct ctt acc atc aag cgt ttc ccc gtc gat ttg tat cac atc aac gat     3704
Ala Leu Thr Ile Lys Arg Phe Pro Val Asp Leu Tyr His Ile Asn Asp
        1205                1210                1215

tac cac gga gcc gtt gca cca gtc cat ctt ctt cca gaa acg att cca     3752
Tyr His Gly Ala Val Ala Pro Val His Leu Leu Pro Glu Thr Ile Pro
    1220                1225                1230

tgt tgt ttg tcc ctc cac aac gcc gag ttc caa ggg tta tgg ccc atg     3800
Cys Cys Leu Ser Leu His Asn Ala Glu Phe Gln Gly Leu Trp Pro Met
1235                1240                1245                1250

cgt aac cag aaa gag aga agt gaa gtc tgt caa gtc tac aac att cca     3848
Arg Asn Gln Lys Glu Arg Ser Glu Val Cys Gln Val Tyr Asn Ile Pro
                1255                1260                1265

gaa gaa gtc gcc atc aaa tac gtt caa ttc gga gat gtg ttc aat ttg     3896
Glu Glu Val Ala Ile Lys Tyr Val Gln Phe Gly Asp Val Phe Asn Leu
            1270                1275                1280

ctc cac gcc gga agt tct tat ctt cgt att cat caa aag ggt ttc ggt     3944
Leu His Ala Gly Ser Ser Tyr Leu Arg Ile His Gln Lys Gly Phe Gly
```

```
        1285                1290                1295

gct gtt ggt gtc agt aag aaa tac gga aag cgt tcc tgg gca aga tat     3992
Ala Val Gly Val Ser Lys Lys Tyr Gly Lys Arg Ser Trp Ala Arg Tyr
    1300                1305                1310

cct att ttc tgg gga ttg agc aag atc ggc aat ctt cct aat cct gat     4040
Pro Ile Phe Trp Gly Leu Ser Lys Ile Gly Asn Leu Pro Asn Pro Asp
1315                1320                1325                1330

cct tct gat act gga gac tgg aat ggt act caa gat tca acg ggc aag     4088
Pro Ser Asp Thr Gly Asp Trp Asn Gly Thr Gln Asp Ser Thr Gly Lys
                1335                1340                1345

aca att gtt aac cag gat ttt gaa agt caa aga gcc gac ctc aaa aga     4136
Thr Ile Val Asn Gln Asp Phe Glu Ser Gln Arg Ala Asp Leu Lys Arg
            1350                1355                1360

caa gca cag gaa tgg gcc ggt ctc gaa caa aat cct aaa gcc gaa ctt     4184
Gln Ala Gln Glu Trp Ala Gly Leu Glu Gln Asn Pro Lys Ala Glu Leu
        1365                1370                1375

ttg gtt ttc gtc ggt aga tgg tcg atg caa aag ggt att gat ttg att     4232
Leu Val Phe Val Gly Arg Trp Ser Met Gln Lys Gly Ile Asp Leu Ile
    1380                1385                1390

gct gat gtt atg cct gct gtt ttg gag gag aat cca aat gtt caa ctt     4280
Ala Asp Val Met Pro Ala Val Leu Glu Glu Asn Pro Asn Val Gln Leu
1395                1400                1405                1410

att tgc gtt ggt cct gtt att gat ctt tat gga aaa ttc gcc gct ctc     4328
Ile Cys Val Gly Pro Val Ile Asp Leu Tyr Gly Lys Phe Ala Ala Leu
                1415                1420                1425

aaa ctc gat gtc atg atg aca aaa tac aag ggc cgt gtc ttc agt aag     4376
Lys Leu Asp Val Met Met Thr Lys Tyr Lys Gly Arg Val Phe Ser Lys
            1430                1435                1440

cct gag ttc act gct ctt cca cca ttc att ttc tcg gga gcc gaa ttc     4424
Pro Glu Phe Thr Ala Leu Pro Pro Phe Ile Phe Ser Gly Ala Glu Phe
        1445                1450                1455

gcc ttg att cca tct aga gat gaa cct ttc ggt ctt gtc gct gta gag     4472
Ala Leu Ile Pro Ser Arg Asp Glu Pro Phe Gly Leu Val Ala Val Glu
    1460                1465                1470

ttt ggt aga aaa gga gct ttg ggt att ggt gct cgt gtt ggt ggt ctt     4520
Phe Gly Arg Lys Gly Ala Leu Gly Ile Gly Ala Arg Val Gly Gly Leu
1475                1480                1485                1490

ggt caa atg cct ggt tgg tgg ttc act gtc gaa tct act acg aca acc     4568
Gly Gln Met Pro Gly Trp Trp Phe Thr Val Glu Ser Thr Thr Thr Thr
                1495                1500                1505

cat atg ctt cat caa ttc aaa caa gcc att aag gaa gcc ttg gcc tcc     4616
His Met Leu His Gln Phe Lys Gln Ala Ile Lys Glu Ala Leu Ala Ser
            1510                1515                1520

aaa ttg gaa act cgt caa ata atg aga gca cga tct gca aag caa cgt     4664
Lys Leu Glu Thr Arg Gln Ile Met Arg Ala Arg Ser Ala Lys Gln Arg
        1525                1530                1535

ttc cca gtc gct caa tgg gtc aat gat ctt gag ata ttg caa agt aaa     4712
Phe Pro Val Ala Gln Trp Val Asn Asp Leu Glu Ile Leu Gln Ser Lys
    1540                1545                1550

gca atc aag att cac gat aag gaa gag gcc aaa cat tct ggt aga cca     4760
Ala Ile Lys Ile His Asp Lys Glu Glu Ala Lys His Ser Gly Arg Pro
```

```
1555                1560                1565                1570

aga tct aga gga aga gat gga gga aat att ttc tat tct tcc gct aag     4808
Arg Ser Arg Gly Arg Asp Gly Gly Asn Ile Phe Tyr Ser Ser Ala Lys
                1575                1580                1585

aga tta tcg agt ttt ggt acc agt ttt gct agt acc tca caa ctt gat     4856
Arg Leu Ser Ser Phe Gly Thr Ser Phe Ala Ser Thr Ser Gln Leu Asp
            1590                1595                1600

gtt gcc agt acc att ggt ggc tct acc aga cca tca tcg cca act cgt     4904
Val Ala Ser Thr Ile Gly Gly Ser Thr Arg Pro Ser Ser Pro Thr Arg
        1605                1610                1615

gtt gaa cag cca ccg gca gga ggt gga ctg agc cga aag ctt tct ttg     4952
Val Glu Gln Pro Pro Ala Gly Gly Gly Leu Ser Arg Lys Leu Ser Leu
    1620                1625                1630

gga tac aga gct ggc ccg gga cat cgc gta acg aaa aat aga ttc aat     5000
Gly Tyr Arg Ala Gly Pro Gly His Arg Val Thr Lys Asn Arg Phe Asn
1635                1640                1645                1650

ggc agt aat gca ccg ggt gct gga atc aaa aat gac gaa gga ctt acc     5048
Gly Ser Asn Ala Pro Gly Ala Gly Ile Lys Asn Asp Glu Gly Leu Thr
                1655                1660                1665

gat gtg gat gag gac agc gat gat gaa gat gcc cga att cgt gac gac     5096
Asp Val Asp Glu Asp Ser Asp Asp Glu Asp Ala Arg Ile Arg Asp Asp
            1670                1675                1680

ata tac gat gaa tac gta ctc acc cct gaa gaa cta gaa gca cag aga     5144
Ile Tyr Asp Glu Tyr Val Leu Thr Pro Glu Glu Leu Glu Ala Gln Arg
        1685                1690                1695

caa caa gag caa tgg aga gga caa gct gaa agt ccc gca gca gga ggc     5192
Gln Gln Glu Gln Trp Arg Gly Gln Ala Glu Ser Pro Ala Ala Gly Gly
    1700                1705                1710

gca ttc ctc gca cca caa cca gga ttc ttg gat tct cca cgc aga agt     5240
Ala Phe Leu Ala Pro Gln Pro Gly Phe Leu Asp Ser Pro Arg Arg Ser
1715                1720                1725                1730

cca agt gcc aat act ggt agt ttg cat atc aac gat tcc gaa ctg tcg     5288
Pro Ser Ala Asn Thr Gly Ser Leu His Ile Asn Asp Ser Glu Leu Ser
                1735                1740                1745

ttg cca cgt ctc aag gat ggc tca aat att tat tcc acc cac tcg ata     5336
Leu Pro Arg Leu Lys Asp Gly Ser Asn Ile Tyr Ser Thr His Ser Ile
            1750                1755                1760

cct ggc tct cct ggt tta gag tca cct gga act ccg gga gca cac gat     5384
Pro Gly Ser Pro Gly Leu Glu Ser Pro Gly Thr Pro Gly Ala His Asp
        1765                1770                1775

tct ttg ctt cct cct cca tca atc ttt gga gcc ggt ggt gac aac gca     5432
Ser Leu Leu Pro Pro Pro Ser Ile Phe Gly Ala Gly Gly Asp Asn Ala
    1780                1785                1790

aac aga gct tcc atg ctt agt ttg agt tcc gtt atc gga gac aag aca     5480
Asn Arg Ala Ser Met Leu Ser Leu Ser Ser Val Ile Gly Asp Lys Thr
1795                1800                1805                1810

gat ttc aaa ctt caa aag gtc gat cca ttc ttc act gat act caa gga     5528
Asp Phe Lys Leu Gln Lys Val Asp Pro Phe Phe Thr Asp Thr Gln Gly
                1815                1820                1825

gaa tac tat caa caa ttt gcc agg aat ctt gat aat ctg gat gga aag     5576
Glu Tyr Tyr Gln Gln Phe Ala Arg Asn Leu Asp Asn Leu Asp Gly Lys
```

```
            1830                1835                1840

aac agt gaa gga caa ttg tgt atc gag gaa tat ctg atc aaa tct gaa     5624
Asn Ser Glu Gly Gln Leu Cys Ile Glu Glu Tyr Leu Ile Lys Ser Glu
        1845                1850                1855

aag aaa tgg ttt gac agg ttc cgt gat gca aaa ttg gga aga aac tca     5672
Lys Lys Trp Phe Asp Arg Phe Arg Asp Ala Lys Leu Gly Arg Asn Ser
    1860                1865                1870

gtt gca cct tcg ttc tct gcc ctt gga agt gga cga aga tct cac gat     5720
Val Ala Pro Ser Phe Ser Ala Leu Gly Ser Gly Arg Arg Ser His Asp
1875                1880                1885                1890

agc cgt gga cga tcc aac agt cgg gat agc ttt acc aac gat acc att     5768
Ser Arg Gly Arg Ser Asn Ser Arg Asp Ser Phe Thr Asn Asp Thr Ile
                1895                1900                1905

acc cac tca agg caa aat tct agt caa gat aac atg gca ctc act ggc     5816
Thr His Ser Arg Gln Asn Ser Ser Gln Asp Asn Met Ala Leu Thr Gly
            1910                1915                1920

gga tct act gcc aat gat caa ttc tta ctt ggc aaa gat tat aaa cca     5864
Gly Ser Thr Ala Asn Asp Gln Phe Leu Leu Gly Lys Asp Tyr Lys Pro
        1925                1930                1935

cct act ggt ctt cgt ctg gtc atg caa tac cgc ttc ttt gat tgg cct     5912
Pro Thr Gly Leu Arg Leu Val Met Gln Tyr Arg Phe Phe Asp Trp Pro
    1940                1945                1950

gtc tac tca ttc tta ctc gcc ttc ggt caa atc atc gca gct aat tct     5960
Val Tyr Ser Phe Leu Leu Ala Phe Gly Gln Ile Ile Ala Ala Asn Ser
1955                1960                1965                1970

tat caa atc act ctt ctt acc ggt act gtg ggt caa agt gca gag aaa     6008
Tyr Gln Ile Thr Leu Leu Thr Gly Thr Val Gly Gln Ser Ala Glu Lys
                1975                1980                1985

ctt tac atc aca gca acg atc tat ctt gtt acc tcg atc atg tgg tgg     6056
Leu Tyr Ile Thr Ala Thr Ile Tyr Leu Val Thr Ser Ile Met Trp Trp
            1990                1995                2000

gaa gtc ttt tcc ttc ttc aaa tca atc tac gtc ctc agt att cca ttc     6104
Glu Val Phe Ser Phe Phe Lys Ser Ile Tyr Val Leu Ser Ile Pro Phe
        2005                2010                2015

atc ttt tac gga ttg gcc ttt ttc ctc ctc ggc atg gct cct ttc gta     6152
Ile Phe Tyr Gly Leu Ala Phe Phe Leu Leu Gly Met Ala Pro Phe Val
    2020                2025                2030

gac ggt gct gat gga agg ggt tgg gtc cag aac gta gct act ggt ttg     6200
Asp Gly Ala Asp Gly Arg Gly Trp Val Gln Asn Val Ala Thr Gly Leu
2035                2040                2045                2050

tat gct gtt gcc gct tca tcg ggt agt att ttc ttc gct ttg aat ttc     6248
Tyr Ala Val Ala Ala Ser Ser Gly Ser Ile Phe Phe Ala Leu Asn Phe
                2055                2060                2065

gga gat gaa gga gga tct cct gtg aaa gcc tgg gtt ttc cgc gcc tgc     6296
Gly Asp Glu Gly Gly Ser Pro Val Lys Ala Trp Val Phe Arg Ala Cys
            2070                2075                2080

atc att caa gga aca caa caa gct tat gtc gtc att ctc tgg ttc tgg     6344
Ile Ile Gln Gly Thr Gln Gln Ala Tyr Val Val Ile Leu Trp Phe Trp
        2085                2090                2095

gga tct tat ctc aac aga cgt caa gct gat ggt ctt gga act tcc ttc     6392
Gly Ser Tyr Leu Asn Arg Arg Gln Ala Asp Gly Leu Gly Thr Ser Phe
```

```
    2100                2105                2110

aca tct gct agt ccc aaa atg tct gca gtc gtt ctt cct att gcc gta     6440
Thr Ser Ala Ser Pro Lys Met Ser Ala Val Val Leu Pro Ile Ala Val
2115                2120                2125                2130

ctc atg tgg gcg gtt ggt ctc atc act ttc ctc gga ttg cca gaa tat     6488
Leu Met Trp Ala Val Gly Leu Ile Thr Phe Leu Gly Leu Pro Glu Tyr
                2135                2140                2145

tat cga caa gcg cct gga aaa gta cca tct ttc tac gcc tcg ctc ttc     6536
Tyr Arg Gln Ala Pro Gly Lys Val Pro Ser Phe Tyr Ala Ser Leu Phe
            2150                2155                2160

aga cgt aag att gtc atg tgg ttt ttc gtc act gtt ttg att cag aac     6584
Arg Arg Lys Ile Val Met Trp Phe Phe Val Thr Val Leu Ile Gln Asn
        2165                2170                2175

ttc ttc ctc tcg acg ctc tat ggt cgt aac tgg ctt tat cta ttc tcc     6632
Phe Phe Leu Ser Thr Leu Tyr Gly Arg Asn Trp Leu Tyr Leu Phe Ser
    2180                2185                2190

tct aac cat gct cct aca tgg gcc atc atc ctt ctc atc atc ttc ttc     6680
Ser Asn His Ala Pro Thr Trp Ala Ile Ile Leu Leu Ile Ile Phe Phe
2195                2200                2205                2210

ttt gga ttc tta tgg gca ggc ttc ctc tgg ttc ttc gca atc ttg agt     6728
Phe Gly Phe Leu Trp Ala Gly Phe Leu Trp Phe Phe Ala Ile Leu Ser
                2215                2220                2225

aaa gat cac agt tgg gtt cta cca att ttc gcc att ggt ctc gga gct     6776
Lys Asp His Ser Trp Val Leu Pro Ile Phe Ala Ile Gly Leu Gly Ala
            2230                2235                2240

cct aga tgg tgt caa atg ctc tgg tct tgc aca ccg ata gga caa tac     6824
Pro Arg Trp Cys Gln Met Leu Trp Ser Cys Thr Pro Ile Gly Gln Tyr
        2245                2250                2255

gta cca tgg gtt ggc ggt ggg ctt gcc agt gct atc ttc tct cga tca     6872
Val Pro Trp Val Gly Gly Gly Leu Ala Ser Ala Ile Phe Ser Arg Ser
    2260                2265                2270

ttg tgg ctt tgg ctc ggt atg ctc gat gca ctt caa ggt gtg gga ttc     6920
Leu Trp Leu Trp Leu Gly Met Leu Asp Ala Leu Gln Gly Val Gly Phe
2275                2280                2285                2290

ggt atg ctg ttg ctt caa act ctt act cgt gtt cat att gct ttc act     6968
Gly Met Leu Leu Leu Gln Thr Leu Thr Arg Val His Ile Ala Phe Thr
                2295                2300                2305

ctc atc gcg gct caa gtt ttg ggt tca ata gct act atg gta gcc aga     7016
Leu Ile Ala Ala Gln Val Leu Gly Ser Ile Ala Thr Met Val Ala Arg
            2310                2315                2320

gca tgt gca cca aat aaa att gga cct gga act gtc ttc cct gat ttt     7064
Ala Cys Ala Pro Asn Lys Ile Gly Pro Gly Thr Val Phe Pro Asp Phe
        2325                2330                2335

agt gac ggg gcg gct gga ctt gga aat ggg tgg ttt tgg atc ggg ttg     7112
Ser Asp Gly Ala Ala Gly Leu Gly Asn Gly Trp Phe Trp Ile Gly Leu
    2340                2345                2350

att tta caa ctc gtt att tgc ggt gga ttt ttc aca ttc ttc aga aaa     7160
Ile Leu Gln Leu Val Ile Cys Gly Gly Phe Phe Thr Phe Phe Arg Lys
2355                2360                2365                2370

gag caa ctt agc aag cct tag                                        7181
Glu Gln Leu Ser Lys Pro  *
```

2375

<210> 476
<211> 911
<212> PRT
<213> Fungi

<400> 476

```
Met Gln Thr Gln Leu Arg His Gly Gly Asp Leu Gln Gly Leu Ile Asp
 1               5                  10                  15
Ser Leu Asp Tyr Ile Gln Gly Met Gly Ile Lys Gly Ile Tyr Ile Ala
            20                  25                  30
Gly Ser Pro Met Ile Asn Ala Pro Trp Gly Ala Asp Gln Tyr Ser Pro
        35                  40                  45
Leu Asp Phe Ser Leu Leu Asp Gln His Phe Gly Asp Ile Asp Leu Trp
    50                  55                  60
Arg Lys Thr Thr Thr Ala Ile His Asp Arg Gly Met Tyr Val Met Leu
65                  70                  75                  80
Asp Asn Thr Met Gly Thr Met Gly Asp Leu Ile Ala Phe Lys Gly Tyr
                85                  90                  95
Glu Asn Ser Ser Thr Pro Phe Thr Leu Thr Glu His Glu Val Leu Trp
            100                 105                 110
Arg Asn Asp Asn Arg Gln Tyr Leu Asp Phe Ser Phe Gly Ser Lys Tyr
        115                 120                 125
Asn Glu Thr Cys Asn Tyr Pro Arg Phe Trp Leu Asp Thr Gly Tyr Thr
    130                 135                 140
Val Lys Glu Asp Val Thr Asp Leu Met Asn Gly Cys Tyr Asp Ser Glu
145                 150                 155                 160
Phe Asp Gln Tyr Gly Asp Thr Glu Ala Phe Gly Val Phe Pro Asp Trp
                165                 170                 175
Gln Arg Gln Leu Ser Lys Phe Ala Ser Val Gln Asp Arg Leu Arg Glu
            180                 185                 190
Trp Val Pro Ser Val Arg Thr Lys Ile Glu Ile Phe Thr Cys Met Ile
        195                 200                 205
Ile Gln Gln Leu Asp Ile Asp Gly Leu Arg Val Asp Lys Ala Thr Gln
    210                 215                 220
Ile Thr Val Asp Ala Leu Gly Ser Phe Ser His Ala Ile Arg Glu Cys
225                 230                 235                 240
Ala Arg Asp Val Asn Lys Thr Asn Phe Met Val Thr Gly Glu Ile Thr
                245                 250                 255
Gly Gly Asn Thr Phe Gly Ser Ile Tyr Leu Gly Arg Gly Arg Gln Pro
            260                 265                 270
Asp Met Leu Pro Ser Ser Met Ala Asp Ala Val Lys Leu Thr Ser Asn
        275                 280                 285
Ser Ser Leu Ala Tyr Phe Ile Arg Asp His Asp Gln Asn Ala Leu Asp
    290                 295                 300
Ser Ala Ala Phe His Tyr Ser Val Tyr Arg Thr Leu Thr Arg Phe Leu
305                 310                 315                 320
Gly Met Asp Gly Asn Leu Ala Ser Gly Tyr Asp Thr Pro Thr Asn Trp
                325                 330                 335
Val Asp Ala Trp Asn Glu Met Leu Val Thr Asn Asp Phe Val Asn Pro
            340                 345                 350
Asn Thr Gly Val Phe Asp Pro Arg His Met Tyr Gly Val Ser Asn Gln
        355                 360                 365
Asp Val Phe Arg Trp Pro Ala Ile Gln Asn Gly Thr Gln Lys Gln Leu
    370                 375                 380
Leu Gly Phe Phe Ile Thr Thr Ile His Met Pro Gly Ile Pro Leu Leu
385                 390                 395                 400
Leu Trp Gly Glu Glu Gln Ala Phe Tyr Val Leu Asp Asn Thr Ala Asp
                405                 410                 415
Asn Tyr Leu Phe Gly Arg Gln Ala Met Ser Pro Ser Gln Ala Trp Gln
            420                 425                 430
Thr His Gly Cys Tyr Ala Leu Gly Ser Gly Thr Tyr Tyr Gln Phe Pro
        435                 440                 445
Asn Asp Ser Cys Thr Arg Gly Cys Asn Asp Asp Trp Asn Ser Arg Asp
    450                 455                 460
His Arg Asp Pro Ser His Pro Val Arg Asn Ile Val Lys Ala Met Tyr
```

```
465                 470                 475                 480
Gln Met Arg Asp Asn Tyr Pro Val Leu Thr Asp Gly Ala Phe Leu Arg
                485                 490                 495
Gln Leu Ser Asn Gln Thr Ile Glu Ile Leu Leu Pro Gly Ser Asn Gly
            500                 505                 510
Thr Ala Thr Glu Thr Gly Met Trp Ser Thr Met Arg Gly Gln Tyr Asp
        515                 520                 525
Gly Ile Gln Asp Leu Ser Lys Ala Gly Gly Gln Ala Asn Gln Ser Val
    530                 535                 540
Trp Met Val Tyr His Asn Asn Tyr Asp Thr Val Thr Tyr Ser Phe Asp
545                 550                 555                 560
Cys Ser Ser Asn Glu Thr Ser Leu Val Ala Pro Phe Leu Asp Gly Val
                565                 570                 575
Thr Val Lys Asn Leu Phe Tyr Pro Tyr Glu Glu Val Thr Leu Lys Ala
            580                 585                 590
Gly Pro Phe Gln Leu Gly Ile Ala Asn Gln Thr Gly Phe Asn Gly Cys
        595                 600                 605
Leu Asp Ser Leu Thr Leu Val Pro Tyr Glu Tyr Lys Ala Tyr Val Pro
    610                 615                 620
Lys Lys Ser Trp Val Gly Pro Ser Pro Met Ile Thr Ser Phe Ser Pro
625                 630                 635                 640
Gly His Asp Ala Arg Leu Val Ser Thr Val Ala Pro Gly Glu Gln Glu
                645                 650                 655
Thr Ile Glu Ile Glu Phe Gly Phe Ser Ala Glu Met Lys Cys Ser Asp
            660                 665                 670
Leu Ile Ser Lys Met Val Ile Thr Ser Ala Thr Glu Asp Gly Gln Val
        675                 680                 685
Ala Thr Leu Ala Asn Ala Thr Gly Glu Tyr Ser Cys Thr Leu Leu Pro
    690                 695                 700
Ser Ala Val Asn Val Thr Glu Phe Val Gly Ser Ile Pro Thr Met Trp
705                 710                 715                 720
Thr Phe Lys Gly Asn Leu Thr Asn Val Ser Asn Gly Val His Ser Ile
                725                 730                 735
Thr Leu Ala Asn Ile Thr Ser Asn Ser Gly Asn Ser Ser Thr Gly Ala
            740                 745                 750
Thr Asp Thr Phe Leu Phe Arg Val Gly Gln Val Asp Asn Pro Ile Ala
        755                 760                 765
Phe Pro Gly Ser Ala Asn Phe Thr Arg Glu Leu Leu His Lys Tyr Asp
    770                 775                 780
Asn Gly Thr Leu Tyr Val Ser His Lys Ala Ala Gly Ala Asp Arg Phe
785                 790                 795                 800
Lys Tyr Ser Leu Asn Trp Gly Ser Ser Trp Ser Glu Trp Leu Pro Tyr
                805                 810                 815
Val Gly Gly Asn Thr Thr Leu Glu Pro Gln Ala Trp Thr Gly Thr Ser
            820                 825                 830
Lys Gln Lys Trp Asp Gly Asp His Val Ile Met Gln Tyr Trp Ser Lys
        835                 840                 845
Leu Thr Gly Ser Ser Asp Val Val Gln His Ala Asp Leu Glu Thr Asn
    850                 855                 860
Lys Trp Ser Arg Arg Leu Pro His Leu Phe Ala Asn Gly Pro Tyr Asn
865                 870                 875                 880
Gln Tyr Gly Tyr Asp Thr Gly Ile Asp Asn Gly Phe Thr Leu Asn Lys
                885                 890                 895
Lys Asp Ser Leu Trp Glu Tyr Lys Phe Met Thr Glu Trp Pro Ala
            900                 905                 910

<210> 477
<211> 1465
<212> PRT
<213> Fungi

<400> 477
Val Met Gln Val Asn Val Trp Gly Met Asn Pro Asp Gly Gln Pro Asp
 1               5                  10                  15
Glu Thr Ile Val Leu Gly Asp Ile Asp Asn Asp Asn Ile Leu Asp Arg
            20                  25                  30
Met Pro Pro Ser Ser Leu Ser Thr Ala Leu Ile Asn Ile Thr Arg Val
        35                  40                  45
```

```
Pro Pro Ser Pro Tyr Leu Gly Tyr Lys Ile Ala Ile Asp Asp Gly Asn
    50                  55                  60
Tyr Arg Tyr Tyr Gln Ile Pro Tyr Gly Asn Arg Ser Tyr Gln Met Leu
65                  70                  75                  80
Met Val Val Leu Phe Trp Ile Val Pro Val Leu Ser Gly Gly Leu Ser
                85                  90                  95
Ile Trp Leu Phe Met Lys Ser Phe Tyr Ala Val Lys Leu Asn Ser Val
            100                 105                 110
Gly Val Lys Glu Ala Lys Asn Tyr Ile Pro Leu Ala Leu Arg Arg Lys
        115                 120                 125
Leu Lys Arg Glu Lys Lys Glu Asp Pro Ser Glu Lys Glu Met Thr Val
    130                 135                 140
Met Gly Gly Leu His Ser Met Phe His Glu Asn Ala Thr Gln Leu Gly
145                 150                 155                 160
Leu Gln Ala Asp Val Ala Ala Thr Gly Arg Arg Thr Val Leu Ile Ala
                165                 170                 175
Thr Met Glu Tyr Asp Ile Glu Asp Trp Ala Ile Lys Ile Lys Ile Gly
            180                 185                 190
Gly Leu Gly Val Met Ala Gln Leu Met Gly Lys Asn Leu Gly His Gln
        195                 200                 205
Asp Leu Ile Trp Val Val Pro Cys Val Gly Gly Ile Asp Tyr Pro Ile
    210                 215                 220
Asp His Gln Ala Glu Pro Met Asp Val Thr Ile Leu Gly Asn Ser Tyr
225                 230                 235                 240
Thr Val Gln Val Gln Tyr His Thr Leu Arg Asn Ile Thr Tyr Val Leu
                245                 250                 255
Leu Asp Ala Pro Val Phe Arg Ala Gln Ser Lys Ser Glu Pro Tyr Pro
            260                 265                 270
Ala Arg Met Asp Asp Leu Asp Ser Ala Val Tyr Tyr Ser Ala Trp Asn
        275                 280                 285
Gln Cys Ile Ala Leu Thr Ile Lys Arg Phe Pro Val Asp Leu Tyr His
    290                 295                 300
Ile Asn Asp Tyr His Gly Ala Val Ala Pro Val His Leu Leu Pro Glu
305                 310                 315                 320
Thr Ile Pro Cys Cys Leu Ser Leu His Asn Ala Glu Phe Gln Gly Leu
                325                 330                 335
Trp Pro Met Arg Asn Gln Lys Glu Arg Ser Glu Val Cys Gln Val Tyr
            340                 345                 350
Asn Ile Pro Glu Glu Val Ala Ile Lys Tyr Val Gln Phe Gly Asp Val
        355                 360                 365
Phe Asn Leu Leu His Ala Gly Ser Ser Tyr Leu Arg Ile His Gln Lys
    370                 375                 380
Gly Phe Gly Ala Val Gly Val Ser Lys Lys Tyr Gly Lys Arg Ser Trp
385                 390                 395                 400
Ala Arg Tyr Pro Ile Phe Trp Gly Leu Ser Lys Ile Gly Asn Leu Pro
                405                 410                 415
Asn Pro Asp Pro Ser Asp Thr Gly Asp Trp Asn Gly Thr Gln Asp Ser
            420                 425                 430
Thr Gly Lys Thr Ile Val Asn Gln Asp Phe Glu Ser Gln Arg Ala Asp
        435                 440                 445
Leu Lys Arg Gln Ala Gln Glu Trp Ala Gly Leu Glu Gln Asn Pro Lys
    450                 455                 460
Ala Glu Leu Leu Val Phe Val Gly Arg Trp Ser Met Gln Lys Gly Ile
465                 470                 475                 480
Asp Leu Ile Ala Asp Val Met Pro Ala Val Leu Glu Glu Asn Pro Asn
                485                 490                 495
Val Gln Leu Ile Cys Val Gly Pro Val Ile Asp Leu Tyr Gly Lys Phe
            500                 505                 510
Ala Ala Leu Lys Leu Asp Val Met Thr Lys Tyr Lys Gly Arg Val
        515                 520                 525
Phe Ser Lys Pro Glu Phe Thr Ala Leu Pro Pro Phe Ile Phe Ser Gly
    530                 535                 540
Ala Glu Phe Ala Leu Ile Pro Ser Arg Asp Glu Pro Phe Gly Leu Val
545                 550                 555                 560
Ala Val Glu Phe Gly Arg Lys Gly Ala Leu Gly Ile Gly Ala Arg Val
                565                 570                 575
Gly Gly Leu Gly Gln Met Pro Gly Trp Trp Phe Thr Val Glu Ser Thr
            580                 585                 590
```

```
Thr Thr Thr His Met Leu His Gln Phe Lys Gln Ala Ile Lys Glu Ala
        595                 600                 605
Leu Ala Ser Lys Leu Glu Thr Arg Gln Ile Met Arg Ala Arg Ser Ala
    610                 615                 620
Lys Gln Arg Phe Pro Val Ala Gln Trp Val Asn Asp Leu Glu Ile Leu
625                 630                 635                 640
Gln Ser Lys Ala Ile Lys Ile His Asp Lys Glu Glu Ala Lys His Ser
                645                 650                 655
Gly Arg Pro Arg Ser Arg Gly Arg Asp Gly Gly Asn Ile Phe Tyr Ser
            660                 665                 670
Ser Ala Lys Arg Leu Ser Ser Phe Gly Thr Ser Phe Ala Ser Thr Ser
        675                 680                 685
Gln Leu Asp Val Ala Ser Thr Ile Gly Gly Ser Thr Arg Pro Ser Ser
    690                 695                 700
Pro Thr Arg Val Glu Gln Pro Pro Ala Gly Gly Gly Leu Ser Arg Lys
705                 710                 715                 720
Leu Ser Leu Gly Tyr Arg Ala Gly Pro Gly His Arg Val Thr Lys Asn
                725                 730                 735
Arg Phe Asn Gly Ser Asn Ala Pro Gly Ala Gly Ile Lys Asn Asp Glu
            740                 745                 750
Gly Leu Thr Asp Val Asp Glu Asp Ser Asp Asp Glu Asp Ala Arg Ile
        755                 760                 765
Arg Asp Asp Ile Tyr Asp Glu Tyr Val Leu Thr Pro Glu Glu Leu Glu
    770                 775                 780
Ala Gln Arg Gln Gln Glu Gln Trp Arg Gly Gln Ala Glu Ser Pro Ala
785                 790                 795                 800
Ala Gly Gly Ala Phe Leu Ala Pro Gln Pro Gly Phe Leu Asp Ser Pro
                805                 810                 815
Arg Arg Ser Pro Ser Ala Asn Thr Gly Ser Leu His Ile Asn Asp Ser
            820                 825                 830
Glu Leu Ser Leu Pro Arg Leu Lys Asp Gly Ser Asn Ile Tyr Ser Thr
        835                 840                 845
His Ser Ile Pro Gly Ser Pro Gly Leu Glu Ser Pro Gly Thr Pro Gly
    850                 855                 860
Ala His Asp Ser Leu Leu Pro Pro Pro Ser Ile Phe Gly Ala Gly Gly
865                 870                 875                 880
Asp Asn Ala Asn Arg Ala Ser Met Leu Ser Leu Ser Ser Val Ile Gly
                885                 890                 895
Asp Lys Thr Asp Phe Lys Leu Gln Lys Val Asp Pro Phe Phe Thr Asp
            900                 905                 910
Thr Gln Gly Glu Tyr Tyr Gln Gln Phe Ala Arg Asn Leu Asp Asn Leu
        915                 920                 925
Asp Gly Lys Asn Ser Glu Gly Gln Leu Cys Ile Glu Glu Tyr Leu Ile
    930                 935                 940
Lys Ser Glu Lys Lys Trp Phe Asp Arg Phe Arg Asp Ala Lys Leu Gly
945                 950                 955                 960
Arg Asn Ser Val Ala Pro Ser Phe Ser Ala Leu Gly Ser Gly Arg Arg
                965                 970                 975
Ser His Asp Ser Arg Gly Arg Ser Asn Ser Arg Asp Ser Phe Thr Asn
            980                 985                 990
Asp Thr Ile Thr His Ser Arg Gln Asn Ser Ser Gln Asp Asn Met Ala
        995                 1000                1005
Leu Thr Gly Gly Ser Thr Ala Asn Asp Gln Phe Leu Leu Gly Lys Asp
    1010                1015                1020
Tyr Lys Pro Pro Thr Gly Leu Arg Leu Val Met Gln Tyr Arg Phe Phe
1025                1030                1035                1040
Asp Trp Pro Val Tyr Ser Phe Leu Leu Ala Phe Gly Gln Ile Ile Ala
                1045                1050                1055
Ala Asn Ser Tyr Gln Ile Thr Leu Leu Thr Gly Thr Val Gly Gln Ser
            1060                1065                1070
Ala Glu Lys Leu Tyr Ile Thr Ala Thr Ile Tyr Leu Val Thr Ser Ile
        1075                1080                1085
Met Trp Trp Glu Val Phe Ser Phe Phe Lys Ser Ile Tyr Val Leu Ser
    1090                1095                1100
Ile Pro Phe Ile Phe Tyr Gly Leu Ala Phe Phe Leu Leu Gly Met Ala
1105                1110                1115                1120
Pro Phe Val Asp Gly Ala Asp Gly Arg Gly Trp Val Gln Asn Val Ala
                1125                1130                1135
```

```
Thr Gly Leu Tyr Ala Val Ala Ala Ser Ser Gly Ser Ile Phe Phe Ala
            1140                1145                1150
Leu Asn Phe Gly Asp Glu Gly Gly Ser Pro Val Lys Ala Trp Val Phe
        1155                1160                1165
Arg Ala Cys Ile Ile Gln Gly Thr Gln Gln Ala Tyr Val Val Ile Leu
    1170                1175                1180
Trp Phe Trp Gly Ser Tyr Leu Asn Arg Arg Gln Ala Asp Gly Leu Gly
1185                1190                1195                1200
Thr Ser Phe Thr Ser Ala Ser Pro Lys Met Ser Ala Val Val Leu Pro
                1205                1210                1215
Ile Ala Val Leu Met Trp Ala Val Gly Leu Ile Thr Phe Leu Gly Leu
            1220                1225                1230
Pro Glu Tyr Tyr Arg Gln Ala Pro Gly Lys Val Pro Ser Phe Tyr Ala
        1235                1240                1245
Ser Leu Phe Arg Arg Lys Ile Val Met Trp Phe Phe Val Thr Val Leu
    1250                1255                1260
Ile Gln Asn Phe Phe Leu Ser Thr Leu Tyr Gly Arg Asn Trp Leu Tyr
1265                1270                1275                1280
Leu Phe Ser Ser Asn His Ala Pro Thr Trp Ala Ile Ile Leu Leu Ile
                1285                1290                1295
Ile Phe Phe Phe Gly Phe Leu Trp Ala Gly Phe Leu Trp Phe Phe Ala
            1300                1305                1310
Ile Leu Ser Lys Asp His Ser Trp Val Leu Pro Ile Phe Ala Ile Gly
        1315                1320                1325
Leu Gly Ala Pro Arg Trp Cys Gln Met Leu Trp Ser Cys Thr Pro Ile
    1330                1335                1340
Gly Gln Tyr Val Pro Trp Val Gly Gly Gly Leu Ala Ser Ala Ile Phe
1345                1350                1355                1360
Ser Arg Ser Leu Trp Leu Trp Leu Gly Met Leu Asp Ala Leu Gln Gly
                1365                1370                1375
Val Gly Phe Gly Met Leu Leu Leu Gln Thr Leu Thr Arg Val His Ile
            1380                1385                1390
Ala Phe Thr Leu Ile Ala Ala Gln Val Leu Gly Ser Ile Ala Thr Met
        1395                1400                1405
Val Ala Arg Ala Cys Ala Pro Asn Lys Ile Gly Pro Gly Thr Val Phe
    1410                1415                1420
Pro Asp Phe Ser Asp Gly Ala Ala Gly Leu Gly Asn Gly Trp Phe Trp
1425                1430                1435                1440
Ile Gly Leu Ile Leu Gln Leu Val Ile Cys Gly Gly Phe Phe Thr Phe
                1445                1450                1455
Phe Arg Lys Glu Gln Leu Ser Lys Pro
            1460                1465
```

<210> 478
<211> 7131
<212> DNA
<213> Fungi

<400> 478

```
atgcaaaccc aattacgaca tggaggagat ctacagggtc tcattgatag tctggactat      60
atacaaggaa tgggaataaa gggaatttat attgctggaa gtcctatgat caatgcgcct     120
tggggcgctg atcaatactc tcctctggat ttcagtcttc tcgaccaaca tttcggagat     180
atcgatttgt ggaggaaaac taccactgcc atccatgaca ggggaatgta tgtgatgctc     240
gacaacacca tgggaaccat gggagacttg attgctttca aaggctatga aaattcttcg     300
actccgttca ccctgaccga gcacgaagtc ctttggcgta atgacaaccg ccaatatctg     360
gatttttcct tcgggtcgaa gtacaatgag acctgcaact accctagatt ctggcttgat     420
actggatata ccgtcaagga agatgtcact gacctcatga atgggtgcta tgattccgaa     480
ttcgatcaat atggtgatac cgaagctttt ggtgtcttcc ctgattggca acgtcagttg     540
tcaaagttcg cctctgtcca ggatcgtttg agagaatggg ttccatccgt aagaactaaa     600
atcgagatct tcacttgcat gatcatccaa caattggaca ttgatggact gcgtgtggac     660
aaggcgacac aaattacggt cgatgcattg ggaagtttca gtcacgcaat cagagagtgc     720
gcaagagatg tgaacaagac aaatttcatg gttactggtg aaatcactgg tggaaacact     780
tttggaagta tttacctcgg gcgtggacgt caacctgaca tgttaccgtc cagcatggca     840
gatgcggtta aattgacctc caattctagt ctggcatatt ttatccgtga tcacgatcaa     900
aacgcgctcg actctgctgc ttttcactat agtgtttatc gcacactgac cagatttctg     960
ggaatggacg gaaacttggc aagtggttat gatacaccaa ctaactgggt agatgcgtgg    1020
aacgaaatgc ttgtaacgaa cgattttgtt aaccctaata ctggagtgtt cgatccaaga    1080
catatgtacg gagtcagcaa tcaggatgtt ttcagatggc cagctatcca gaatggcaca    1140
```

caaaaacaac ttctagggtt cttcatcacg actatacaca tgcctggtat tcctttgctt 1200
ttgtggggtg aagaacaagc attctatgtc ttagacaata cagccgacaa ttatctcttc 1260
ggtcgacaag ctatgtctcc ttctcaagca tggcagactc acggatgcta cgcccttggc 1320
agtggtacat attatcaatt ccccaatgac tcttgcactc gaggttgcaa tgatgattgg 1380
aacagccgcg atcacagaga tccatctcat ccagttagaa atattgtcaa ggcaatgtac 1440
cagatgcgag acaactatcc agttttgacc gatggagcat ttttacgaca actttccaac 1500
cagacaatcg aaatcctctt accaggtagt aacggaacag ctactgaaac tggaatgtgg 1560
agtaccatgc gaggtcaata cgatggtatt caagatttat ctaaagctgg tggtcaagca 1620
aaccagtctg tttggatggt ttaccacaac aattatgaca cagttacata ttcttttgat 1680
tgcagtagca acgaaacctc gcttgttgct cctttccttg acggtgtcac tgtcaaaaat 1740
ttgttctatc cgtatgagga ggttacattg aaagcaggtc cttttcagct aggaattgca 1800
aatcagactg gcttcaatgg ctgtctcgat agtctcactc tcgttccata cgagtataaa 1860
gcgtatgtgc caaaaaagag ctgggttgga ccttcgccta tgattacctc attttcacca 1920
ggacacgatg cacgacttgt ttctacagtt gctcccggag aacaagaaac cattgaaatt 1980
gaatttggct tcagtgccga aatgaagtgc tctgatctca tttcgaaaat ggtaattacc 2040
tcggcaacag aagacggaca ggttgcaaca ctagccaatg ccaccggtga atattcttgt 2100
actttgcttc caagtgcggt caatgtaacc gaattcgtcg gatcaattcc aaccatgtgg 2160
actttcaaag gaaatcttac caatgtatcc aacggtgtcc actcaatcac tctagccaat 2220
atcaccagca actctggaaa ttcctcaact ggagccacag atacattctt gttccgtgtt 2280
ggacaagttg ataatcctat cgctttcccc ggttccgcaa acttcacaag agaattactt 2340
cacaagtatg acaatggcac attgtatgtc agccacaaag ctgctggagc agatcggttc 2400
aaatattctt tgaactgggg aagttcatgg tctgaatggt taccatatgt tggtggaaat 2460
accacacttg aaccacaggc ttggactggt acaagcaaac aaaaatggga cggagatcac 2520
gtcatcatgc aatattggag caaacttact ggtagcagtg atgtagttca acatgctgat 2580
ctagaaacca acaaatggtc gcgtcgcctt ccgcacttgt ttgctaatgg tccatacaac 2640
cagtatggat atgatactgg tatcgataat ggtttcactt tgaacaagaa agacagtctc 2700
tgggaatata agtttatgac tgaatggcca gccgtcatgc aagtgaatgt ttggggtatg 2760
aatcctgatg ggcaacctga cgaaactatc gtgttgggag atatcgacaa tgataatatc 2820
cttgatcgta tgccaccttc atcgctatcc acggcgctca tcaacatcac acgcgtgcca 2880
ccatcaccat atctcggata caaaattgca atcgacgacg gtaattatag atactatcaa 2940
atcccttacg gaaatcgttc ctaccaaatg ttgatggttg tccttttctg gattgtccct 3000
gtactcagtg gtggtctcag tatctggctg ttcatgaaat ctttctacgc cgtcaaactc 3060
aacagtgtcg gtgtcaagga ggcaaagaac tatatcccat tagccttgcg cagaaagctg 3120
aagagagaaa agaaagagga cccatctgag aaagaaatga ccgtcatggg aggattgcac 3180
tcaatgttcc acgaaaatgc cacacaactt ggactgcagg ctgatgttgc agcgaccgga 3240
cgacgaactg tacttattgc tactatggaa tatgatattg aagattgggc tatcaagatc 3300
aagattggtg gtcttggtgt catggcacaa ttgatgggaa agaatttggg tcatcaagat 3360
ctcatttggg tcgttccttg tgttggtggt atcgattatc ctatcgatca ccaagctgaa 3420
cctatggatg tcacaattct tggaaattcc tacactgttc aagttcagta ccatactcta 3480
cgaaacatca cttacgttct tctcgatgca ccagttttca gagcacaatc taagtctgag 3540
ccatacccag ctcgtatgga tgatttagac tcggccgttt actacagtgc ttggaatcaa 3600
tgtatcgctc ttaccatcaa gcgtttcccc gtcgatttgt atcacatcaa cgattaccac 3660
ggagccgttg caccagtcca tcttcttcca gaaacgattc catgttgttt gtccctccac 3720
aacgccgagt tccaagggtt atggcccatg cgtaaccaga aagagagaag tgaagtctgt 3780
caagtctaca acattccaga agaagtcgcc atcaaatacg ttcaattcgg agatgtgttc 3840
aatttgctcc acgccggaag ttcttatctt cgtattcatc aaaagggttt cggtgctgtt 3900
ggtgtcagta agaaatacgg aaagcgttcc tgggcaagat atcctatttt ctggggattg 3960
agcaagatcg gcaatcttcc taatcctgat ccttctgata ctggagactg gaatggtact 4020
caagattcaa cgggcaagac aattgttaac caggattttg aaagtcaaag agccgacctc 4080
aaaagacaag cacaggaatg ggccggtctc gaacaaaatc ctaaagccga acttttggtt 4140
ttcgtcggta gatggtcgat gcaaaagggt attgatttga ttgctgatgt tatgcctgct 4200
gttttggagg agaatccaaa tgttcaactt atttgcgttg gtcctgttat tgatctttat 4260
ggaaaattcg ccgctctcaa actcgatgtc atgatgacaa aatacaaggg ccgtgtcttc 4320
agtaagcctg agttcactgc tcttccacca ttcattttct cgggagccga attcgccttg 4380
attccatcta gagatgaacc tttcggtctt gtcgctgtag agtttggtag aaaaggagct 4440
ttgggtattg gtgctcgtgt tggtggtctt ggtcaaatgc ctggttggtg gttcactgtc 4500
gaatctacta cgacaaccca tatgcttcat caattcaaac aagccattaa ggaagccttg 4560
gcctccaaat tggaaactcg tcaaataatg agagcacgat ctgcaaagca acgtttccca 4620
gtcgctcaat gggtcaatga tcttgagata ttgcaaagta aagcaatcaa gattcacgat 4680
aaggaagagg ccaaacattc tggtagacca agatctagag gaagagatgg aggaaatatt 4740
ttctattctt ccgctaagag attatcgagt tttggtacca gttttgctag tacctcacaa 4800
cttgatgttg ccagtaccat tggtggctct accagaccat catcgccaac tcgtgttgaa 4860
cagccaccgg caggaggtgg actgagccga aagctttctt tgggatacag agctggcccg 4920
ggacatcgcg taacgaaaaa tagattcaat ggcagtaatg caccgggtgc tggaatcaaa 4980
aatgacgaag gacttaccga tgtggatgag gacagcgatg atgaagatgc ccgaattcgt 5040
gacgacatat acgatgaata cgtactcacc cctgaagaac tagaagcaca gagacaacaa 5100
gagcaatgga gaggacaagc tgaaagtccc gcagcaggag gcgcattcct cgcaccacaa 5160
ccaggattct tggattctcc acgcagaagt ccaagtgcca atactggtag tttgcatatc 5220

```
aacgattccg aactgtcgtt gccacgtctc aaggatggct caaatattta ttccacccac    5280
tcgatacctg gctctcctgg tttagagtca cctggaactc cgggagcaca cgattctttg    5340
cttcctcctc catcaatctt tggagccggt ggtgacaacg caaacagagc ttccatgctt    5400
agtttgagtt ccgttatcgg agacaagaca gatttcaaac ttcaaaaggt cgatccattc    5460
ttcactgata ctcaaggaga atactatcaa caatttgcca ggaatcttga taatctggat    5520
ggaaagaaca gtgaaggaca attgtgtatc gaggaatatc tgatcaaatc tgaaaagaaa    5580
tggtttgaca ggttccgtga tgcaaaattg ggaagaaact cagttgcacc ttcgttctct    5640
gcccttggaa gtggacgaag atctcacgat agccgtggac gatccaacag tcgggatagc    5700
tttaccaacg ataccattac ccactcaagg caaaattcta gtcaagataa catggcactc    5760
actggcggat ctactgccaa tgatcaattc ttacttggca aagattataa accacctact    5820
ggtcttcgtc tggtcatgca ataccgcttc tttgattggc ctgtctactc attcttactc    5880
gccttcggtc aaatcatcgc agctaattct tatcaaatca ctcttcttac cggtactgtg    5940
ggtcaaagtg cagagaaact ttacatcaca gcaacgatct atcttgttac ctcgatcatg    6000
tggtgggaag tcttttcctt cttcaaatca atctacgtcc tcagtattcc attcatcttt    6060
tacggattgg ccttttttcct cctcggcatg gctcctttcg tagacggtgc tgatggaagg    6120
ggttgggtcc agaacgtagc tactggtttg tatgctgttg ccgcttcatc gggtagtatt    6180
ttcttcgctt tgaatttcgg agatgaagga ggatctcctg tgaaagcctg ggttttccgc    6240
gcctgcatca ttcaaggaac acaacaagct tatgtcgtca ttctctggtt ctggggatct    6300
tatctcaaca gacgtcaagc tgatggtctt ggaacttcct tcacatctgc tagtcccaaa    6360
atgtctgcag tcgttcttcc tattgccgta ctcatgtggg cggttggtct catcactttc    6420
ctcggattgc cagaatatta tcgacaagcg cctggaaaag taccatcttt ctacgcctcg    6480
ctcttcagac gtaagattgt catgtggttt ttcgtcactg ttttgattca gaacttcttc    6540
ctctcgacgc tctatggtcg taactggctt tatctattct cctctaacca tgctcctaca    6600
tgggccatca tccttctcat catcttcttc tttggattct tatgggcagg cttcctctgg    6660
ttcttcgcaa tcttgagtaa agatcacagt tgggttctac caattttcgc cattggtctc    6720
ggagctccta gatggtgtca aatgctctgg tcttgcacac cgataggaca atacgtacca    6780
tgggttggcg gtgggcttgc cagtgctatc ttctctcgat cattgtggct ttggctcggt    6840
atgctcgatg cacttcaagg tgtgggattc ggtatgctgt tgcttcaaac tcttactcgt    6900
gttcatattg ctttcactct catcgcggct caagttttgg gttcaatagc tactatggta    6960
gccagagcat gtgcaccaaa taaaattgga cctggaactg tcttccctga ttttagtgac    7020
ggggcggctg gacttggaaa tgggtggttt tggatcgggt tgattttaca actcgttatt    7080
tgcggtggat ttttcacatt cttcagaaaa gagcaactta gcaagcctta g              7131
```

<210> 479
<211> 2376
<212> PRT
<213> Fungi

<400> 479

```
Met Gln Thr Gln Leu Arg His Gly Gly Asp Leu Gln Gly Leu Ile Asp
 1               5                  10                  15
Ser Leu Asp Tyr Ile Gln Gly Met Gly Ile Lys Gly Ile Tyr Ile Ala
            20                  25                  30
Gly Ser Pro Met Ile Asn Ala Pro Trp Gly Ala Asp Gln Tyr Ser Pro
        35                  40                  45
Leu Asp Phe Ser Leu Leu Asp Gln His Phe Gly Asp Ile Asp Leu Trp
    50                  55                  60
Arg Lys Thr Thr Thr Ala Ile His Asp Arg Gly Met Tyr Val Met Leu
65                  70                  75                  80
Asp Asn Thr Met Gly Thr Met Gly Asp Leu Ile Ala Phe Lys Gly Tyr
                85                  90                  95
Glu Asn Ser Ser Thr Pro Phe Thr Leu Thr Glu His Glu Val Leu Trp
            100                 105                 110
Arg Asn Asp Asn Arg Gln Tyr Leu Asp Phe Ser Phe Gly Ser Lys Tyr
        115                 120                 125
Asn Glu Thr Cys Asn Tyr Pro Arg Phe Trp Leu Asp Thr Gly Tyr Thr
    130                 135                 140
Val Lys Glu Asp Val Thr Asp Leu Met Asn Gly Cys Tyr Asp Ser Glu
145                 150                 155                 160
Phe Asp Gln Tyr Gly Asp Thr Glu Ala Phe Gly Val Phe Pro Asp Trp
                165                 170                 175
Gln Arg Gln Leu Ser Lys Phe Ala Ser Val Gln Asp Arg Leu Arg Glu
            180                 185                 190
Trp Val Pro Ser Val Arg Thr Lys Ile Glu Ile Phe Thr Cys Met Ile
        195                 200                 205
Ile Gln Gln Leu Asp Ile Asp Gly Leu Arg Val Asp Lys Ala Thr Gln
```

210 215 220
Ile Thr Val Asp Ala Leu Gly Ser Phe Ser His Ala Ile Arg Glu Cys
225 230 235 240
Ala Arg Asp Val Asn Lys Thr Asn Phe Met Val Thr Gly Glu Ile Thr
245 250 255
Gly Gly Asn Thr Phe Gly Ser Ile Tyr Leu Gly Arg Gly Arg Gln Pro
260 265 270
Asp Met Leu Pro Ser Ser Met Ala Asp Ala Val Lys Leu Thr Ser Asn
275 280 285
Ser Ser Leu Ala Tyr Phe Ile Arg Asp His Asp Gln Asn Ala Leu Asp
290 295 300
Ser Ala Ala Phe His Tyr Ser Val Tyr Arg Thr Leu Thr Arg Phe Leu
305 310 315 320
Gly Met Asp Gly Asn Leu Ala Ser Gly Tyr Asp Thr Pro Thr Asn Trp
325 330 335
Val Asp Ala Trp Asn Glu Met Leu Val Thr Asn Asp Phe Val Asn Pro
340 345 350
Asn Thr Gly Val Phe Asp Pro Arg His Met Tyr Gly Val Ser Asn Gln
355 360 365
Asp Val Phe Arg Trp Pro Ala Ile Gln Asn Gly Thr Gln Lys Gln Leu
370 375 380
Leu Gly Phe Phe Ile Thr Thr Ile His Met Pro Gly Ile Pro Leu Leu
385 390 395 400
Leu Trp Gly Glu Glu Gln Ala Phe Tyr Val Leu Asp Asn Thr Ala Asp
405 410 415
Asn Tyr Leu Phe Gly Arg Gln Ala Met Ser Pro Ser Gln Ala Trp Gln
420 425 430
Thr His Gly Cys Tyr Ala Leu Gly Ser Gly Thr Tyr Tyr Gln Phe Pro
435 440 445
Asn Asp Ser Cys Thr Arg Gly Cys Asn Asp Asp Trp Asn Ser Arg Asp
450 455 460
His Arg Asp Pro Ser His Pro Val Arg Asn Ile Val Lys Ala Met Tyr
465 470 475 480
Gln Met Arg Asp Asn Tyr Pro Val Leu Thr Asp Gly Ala Phe Leu Arg
485 490 495
Gln Leu Ser Asn Gln Thr Ile Glu Ile Leu Leu Pro Gly Ser Asn Gly
500 505 510
Thr Ala Thr Glu Thr Gly Met Trp Ser Thr Met Arg Gly Gln Tyr Asp
515 520 525
Gly Ile Gln Asp Leu Ser Lys Ala Gly Gly Gln Ala Asn Gln Ser Val
530 535 540
Trp Met Val Tyr His Asn Asn Tyr Asp Thr Val Thr Tyr Ser Phe Asp
545 550 555 560
Cys Ser Ser Asn Glu Thr Ser Leu Val Ala Pro Phe Leu Asp Gly Val
565 570 575
Thr Val Lys Asn Leu Phe Tyr Pro Tyr Glu Glu Val Thr Leu Lys Ala
580 585 590
Gly Pro Phe Gln Leu Gly Ile Ala Asn Gln Thr Gly Phe Asn Gly Cys
595 600 605
Leu Asp Ser Leu Thr Leu Val Pro Tyr Glu Tyr Lys Ala Tyr Val Pro
610 615 620
Lys Lys Ser Trp Val Gly Pro Ser Pro Met Ile Thr Ser Phe Ser Pro
625 630 635 640
Gly His Asp Ala Arg Leu Val Ser Thr Val Ala Pro Gly Glu Gln Glu
645 650 655
Thr Ile Glu Ile Glu Phe Gly Phe Ser Ala Glu Met Lys Cys Ser Asp
660 665 670
Leu Ile Ser Lys Met Val Ile Thr Ser Ala Thr Glu Asp Gly Gln Val
675 680 685
Ala Thr Leu Ala Asn Ala Thr Gly Glu Tyr Ser Cys Thr Leu Leu Pro
690 695 700
Ser Ala Val Asn Val Thr Glu Phe Val Gly Ser Ile Pro Thr Met Trp
705 710 715 720
Thr Phe Lys Gly Asn Leu Thr Asn Val Ser Asn Gly Val His Ser Ile
725 730 735
Thr Leu Ala Asn Ile Thr Ser Asn Ser Gly Asn Ser Ser Thr Gly Ala
740 745 750
Thr Asp Thr Phe Leu Phe Arg Val Gly Gln Val Asp Asn Pro Ile Ala

```
        755                 760                 765
Phe Pro Gly Ser Ala Asn Phe Thr Arg Glu Leu Leu His Lys Tyr Asp
    770                 775                 780
Asn Gly Thr Leu Tyr Val Ser His Lys Ala Ala Gly Ala Asp Arg Phe
785                 790                 795                 800
Lys Tyr Ser Leu Asn Trp Gly Ser Ser Trp Ser Glu Trp Leu Pro Tyr
                805                 810                 815
Val Gly Gly Asn Thr Thr Leu Glu Pro Gln Ala Trp Thr Gly Thr Ser
            820                 825                 830
Lys Gln Lys Trp Asp Gly Asp His Val Ile Met Gln Tyr Trp Ser Lys
        835                 840                 845
Leu Thr Gly Ser Ser Asp Val Val Gln His Ala Asp Leu Glu Thr Asn
    850                 855                 860
Lys Trp Ser Arg Arg Leu Pro His Leu Phe Ala Asn Gly Pro Tyr Asn
865                 870                 875                 880
Gln Tyr Gly Tyr Asp Thr Gly Ile Asp Asn Gly Phe Thr Leu Asn Lys
                885                 890                 895
Lys Asp Ser Leu Trp Glu Tyr Lys Phe Met Thr Glu Trp Pro Ala Val
            900                 905                 910
Met Gln Val Asn Val Trp Gly Met Asn Pro Asp Gly Gln Pro Asp Glu
        915                 920                 925
Thr Ile Val Leu Gly Asp Ile Asp Asn Asp Asn Ile Leu Asp Arg Met
    930                 935                 940
Pro Pro Ser Ser Leu Ser Thr Ala Leu Ile Asn Ile Thr Arg Val Pro
945                 950                 955                 960
Pro Ser Pro Tyr Leu Gly Tyr Lys Ile Ala Ile Asp Asp Gly Asn Tyr
                965                 970                 975
Arg Tyr Tyr Gln Ile Pro Tyr Gly Asn Arg Ser Tyr Gln Met Leu Met
            980                 985                 990
Val Val Leu Phe Trp Ile Val Pro Val Leu Ser Gly Gly Leu Ser Ile
        995                 1000                1005
Trp Leu Phe Met Lys Ser Phe Tyr Ala Val Lys Leu Asn Ser Val Gly
    1010                1015                1020
Val Lys Glu Ala Lys Asn Tyr Ile Pro Leu Ala Leu Arg Arg Lys Leu
1025                1030                1035                1040
Lys Arg Glu Lys Lys Glu Asp Pro Ser Glu Lys Glu Met Thr Val Met
                1045                1050                1055
Gly Gly Leu His Ser Met Phe His Glu Asn Ala Thr Gln Leu Gly Leu
            1060                1065                1070
Gln Ala Asp Val Ala Ala Thr Gly Arg Arg Thr Val Leu Ile Ala Thr
        1075                1080                1085
Met Glu Tyr Asp Ile Glu Asp Trp Ala Ile Lys Ile Lys Ile Gly Gly
    1090                1095                1100
Leu Gly Val Met Ala Gln Leu Met Gly Lys Asn Leu Gly His Gln Asp
1105                1110                1115                1120
Leu Ile Trp Val Val Pro Cys Val Gly Gly Ile Asp Tyr Pro Ile Asp
                1125                1130                1135
His Gln Ala Glu Pro Met Asp Val Thr Ile Leu Gly Asn Ser Tyr Thr
            1140                1145                1150
Val Gln Val Gln Tyr His Thr Leu Arg Asn Ile Thr Tyr Val Leu Leu
        1155                1160                1165
Asp Ala Pro Val Phe Arg Ala Gln Ser Lys Ser Glu Pro Tyr Pro Ala
    1170                1175                1180
Arg Met Asp Asp Leu Asp Ser Ala Val Tyr Tyr Ser Ala Trp Asn Gln
1185                1190                1195                1200
Cys Ile Ala Leu Thr Ile Lys Arg Phe Pro Val Asp Leu Tyr His Ile
                1205                1210                1215
Asn Asp Tyr His Gly Ala Val Ala Pro Val His Leu Leu Pro Glu Thr
            1220                1225                1230
Ile Pro Cys Cys Leu Ser Leu His Asn Ala Glu Phe Gln Gly Leu Trp
        1235                1240                1245
Pro Met Arg Asn Gln Lys Glu Arg Ser Glu Val Cys Gln Val Tyr Asn
    1250                1255                1260
Ile Pro Glu Glu Val Ala Ile Lys Tyr Val Gln Phe Gly Asp Val Phe
1265                1270                1275                1280
Asn Leu Leu His Ala Gly Ser Ser Tyr Leu Arg Ile His Gln Lys Gly
                1285                1290                1295
Phe Gly Ala Val Gly Val Ser Lys Lys Tyr Gly Lys Arg Ser Trp Ala
```

```
            1300                1305                1310
Arg Tyr Pro Ile Phe Trp Gly Leu Ser Lys Ile Gly Asn Leu Pro Asn
        1315                1320                1325
Pro Asp Pro Ser Asp Thr Gly Asp Trp Asn Gly Thr Gln Asp Ser Thr
    1330                1335                1340
Gly Lys Thr Ile Val Asn Gln Asp Phe Glu Ser Gln Arg Ala Asp Leu
1345                1350                1355                1360
Lys Arg Gln Ala Gln Glu Trp Ala Gly Leu Glu Gln Asn Pro Lys Ala
                1365                1370                1375
Glu Leu Leu Val Phe Val Gly Arg Trp Ser Met Gln Lys Gly Ile Asp
            1380                1385                1390
Leu Ile Ala Asp Val Met Pro Ala Val Leu Glu Glu Asn Pro Asn Val
        1395                1400                1405
Gln Leu Ile Cys Val Gly Pro Val Ile Asp Leu Tyr Gly Lys Phe Ala
    1410                1415                1420
Ala Leu Lys Leu Asp Val Met Met Thr Lys Tyr Lys Gly Arg Val Phe
1425                1430                1435                1440
Ser Lys Pro Glu Phe Thr Ala Leu Pro Pro Phe Ile Phe Ser Gly Ala
                1445                1450                1455
Glu Phe Ala Leu Ile Pro Ser Arg Asp Glu Pro Phe Gly Leu Val Ala
            1460                1465                1470
Val Glu Phe Gly Arg Lys Gly Ala Leu Gly Ile Gly Ala Arg Val Gly
        1475                1480                1485
Gly Leu Gly Gln Met Pro Gly Trp Trp Phe Thr Val Glu Ser Thr Thr
    1490                1495                1500
Thr Thr His Met Leu His Gln Phe Lys Gln Ala Ile Lys Glu Ala Leu
1505                1510                1515                1520
Ala Ser Lys Leu Glu Thr Arg Gln Ile Met Arg Ala Arg Ser Ala Lys
                1525                1530                1535
Gln Arg Phe Pro Val Ala Gln Trp Val Asn Asp Leu Glu Ile Leu Gln
            1540                1545                1550
Ser Lys Ala Ile Lys Ile His Asp Lys Glu Glu Ala Lys His Ser Gly
        1555                1560                1565
Arg Pro Arg Ser Arg Gly Arg Asp Gly Gly Asn Ile Phe Tyr Ser Ser
    1570                1575                1580
Ala Lys Arg Leu Ser Ser Phe Gly Thr Ser Phe Ala Ser Thr Ser Gln
1585                1590                1595                1600
Leu Asp Val Ala Ser Thr Ile Gly Gly Ser Thr Arg Pro Ser Ser Pro
                1605                1610                1615
Thr Arg Val Glu Gln Pro Pro Ala Gly Gly Gly Leu Ser Arg Lys Leu
            1620                1625                1630
Ser Leu Gly Tyr Arg Ala Gly Pro Gly His Arg Val Thr Lys Asn Arg
        1635                1640                1645
Phe Asn Gly Ser Asn Ala Pro Gly Ala Gly Ile Lys Asn Asp Glu Gly
    1650                1655                1660
Leu Thr Asp Val Asp Glu Asp Ser Asp Asp Glu Asp Ala Arg Ile Arg
1665                1670                1675                1680
Asp Asp Ile Tyr Asp Glu Tyr Val Leu Thr Pro Glu Glu Leu Glu Ala
                1685                1690                1695
Gln Arg Gln Gln Glu Gln Trp Arg Gly Gln Ala Glu Ser Pro Ala Ala
            1700                1705                1710
Gly Gly Ala Phe Leu Ala Pro Gln Pro Gly Phe Leu Asp Ser Pro Arg
        1715                1720                1725
Arg Ser Pro Ser Ala Asn Thr Gly Ser Leu His Ile Asn Asp Ser Glu
    1730                1735                1740
Leu Ser Leu Pro Arg Leu Lys Asp Gly Ser Asn Ile Tyr Ser Thr His
1745                1750                1755                1760
Ser Ile Pro Gly Ser Pro Gly Leu Glu Ser Pro Gly Thr Pro Gly Ala
                1765                1770                1775
His Asp Ser Leu Leu Pro Pro Pro Ser Ile Phe Gly Ala Gly Gly Asp
            1780                1785                1790
Asn Ala Asn Arg Ala Ser Met Leu Ser Leu Ser Ser Val Ile Gly Asp
        1795                1800                1805
Lys Thr Asp Phe Lys Leu Gln Lys Val Asp Pro Phe Phe Thr Asp Thr
    1810                1815                1820
Gln Gly Glu Tyr Tyr Gln Gln Phe Ala Arg Asn Leu Asp Asn Leu Asp
1825                1830                1835                1840
Gly Lys Asn Ser Glu Gly Gln Leu Cys Ile Glu Glu Tyr Leu Ile Lys
```

```
                1845                1850                1855
Ser Glu Lys Lys Trp Phe Asp Arg Phe Arg Asp Ala Lys Leu Gly Arg
            1860                1865                1870
Asn Ser Val Ala Pro Ser Phe Ser Ala Leu Gly Ser Gly Arg Arg Ser
        1875                1880                1885
His Asp Ser Arg Gly Arg Ser Asn Ser Arg Asp Ser Phe Thr Asn Asp
    1890                1895                1900
Thr Ile Thr His Ser Arg Gln Asn Ser Ser Gln Asp Asn Met Ala Leu
1905                1910                1915                1920
Thr Gly Gly Ser Thr Ala Asn Asp Gln Phe Leu Leu Gly Lys Asp Tyr
                1925                1930                1935
Lys Pro Pro Thr Gly Leu Arg Leu Val Met Gln Tyr Arg Phe Phe Asp
            1940                1945                1950
Trp Pro Val Tyr Ser Phe Leu Leu Ala Phe Gly Gln Ile Ile Ala Ala
        1955                1960                1965
Asn Ser Tyr Gln Ile Thr Leu Leu Thr Gly Thr Val Gly Gln Ser Ala
    1970                1975                1980
Glu Lys Leu Tyr Ile Thr Ala Thr Ile Tyr Leu Val Thr Ser Ile Met
1985                1990                1995                2000
Trp Trp Glu Val Phe Ser Phe Phe Lys Ser Ile Tyr Val Leu Ser Ile
                2005                2010                2015
Pro Phe Ile Phe Tyr Gly Leu Ala Phe Phe Leu Leu Gly Met Ala Pro
            2020                2025                2030
Phe Val Asp Gly Ala Asp Gly Arg Gly Trp Val Gln Asn Val Ala Thr
        2035                2040                2045
Gly Leu Tyr Ala Val Ala Ala Ser Ser Gly Ser Ile Phe Phe Ala Leu
    2050                2055                2060
Asn Phe Gly Asp Glu Gly Gly Ser Pro Val Lys Ala Trp Val Phe Arg
2065                2070                2075                2080
Ala Cys Ile Ile Gln Gly Thr Gln Gln Ala Tyr Val Val Ile Leu Trp
                2085                2090                2095
Phe Trp Gly Ser Tyr Leu Asn Arg Arg Gln Ala Asp Gly Leu Gly Thr
            2100                2105                2110
Ser Phe Thr Ser Ala Ser Pro Lys Met Ser Ala Val Val Leu Pro Ile
        2115                2120                2125
Ala Val Leu Met Trp Ala Val Gly Leu Ile Thr Phe Leu Gly Leu Pro
    2130                2135                2140
Glu Tyr Tyr Arg Gln Ala Pro Gly Lys Val Pro Ser Phe Tyr Ala Ser
2145                2150                2155                2160
Leu Phe Arg Arg Lys Ile Val Met Trp Phe Phe Val Thr Val Leu Ile
                2165                2170                2175
Gln Asn Phe Phe Leu Ser Thr Leu Tyr Gly Arg Asn Trp Leu Tyr Leu
            2180                2185                2190
Phe Ser Ser Asn His Ala Pro Thr Trp Ala Ile Ile Leu Leu Ile Ile
        2195                2200                2205
Phe Phe Phe Gly Phe Leu Trp Ala Gly Phe Leu Trp Phe Phe Ala Ile
    2210                2215                2220
Leu Ser Lys Asp His Ser Trp Val Leu Pro Ile Phe Ala Ile Gly Leu
2225                2230                2235                2240
Gly Ala Pro Arg Trp Cys Gln Met Leu Trp Ser Cys Thr Pro Ile Gly
                2245                2250                2255
Gln Tyr Val Pro Trp Val Gly Gly Gly Leu Ala Ser Ala Ile Phe Ser
            2260                2265                2270
Arg Ser Leu Trp Leu Trp Leu Gly Met Leu Asp Ala Leu Gln Gly Val
        2275                2280                2285
Gly Phe Gly Met Leu Leu Leu Gln Thr Leu Thr Arg Val His Ile Ala
    2290                2295                2300
Phe Thr Leu Ile Ala Ala Gln Val Leu Gly Ser Ile Ala Thr Met Val
2305                2310                2315                2320
Ala Arg Ala Cys Ala Pro Asn Lys Ile Gly Pro Gly Thr Val Phe Pro
                2325                2330                2335
Asp Phe Ser Asp Gly Ala Ala Gly Leu Gly Asn Gly Trp Phe Trp Ile
            2340                2345                2350
Gly Leu Ile Leu Gln Leu Val Ile Cys Gly Gly Phe Phe Thr Phe Phe
        2355                2360                2365
Arg Lys Glu Gln Leu Ser Lys Pro
    2370                2375
```

<210> 480
<211> 1511
<212> DNA
<213> Fungi

<220>
<221> CDS
<222> (1)...(263)
<223> Exon

<221> CDS
<222> (312)...(417)
<223> Exon

<221> CDS
<222> (498)...(1510)
<223> Exon

<400> 480

```
atg aag ttc tct ctt ctt gcg acc atc gtc gct agc att agt ccg ctt       48
Met Lys Phe Ser Leu Leu Ala Thr Ile Val Ala Ser Ile Ser Pro Leu
 1               5                  10                  15

gca cgc gca gca gat gca aac gct tgg aag tcg cga aat att tac ttc       96
Ala Arg Ala Ala Asp Ala Asn Ala Trp Lys Ser Arg Asn Ile Tyr Phe
            20                  25                  30

gca ctt act gat cgt gtt gcg cgg agt gct gat gat aat ggc ggc agt      144
Ala Leu Thr Asp Arg Val Ala Arg Ser Ala Asp Asp Asn Gly Gly Ser
         35                  40                  45

gca tgc gga aac ctc ggg aat tac tgt ggt gga act ttt aag ggc tta      192
Ala Cys Gly Asn Leu Gly Asn Tyr Cys Gly Gly Thr Phe Lys Gly Leu
     50                  55                  60

gag tca aag ctt gat tat atc aag ggc atg gga ttt gat gct att tgg      240
Glu Ser Lys Leu Asp Tyr Ile Lys Gly Met Gly Phe Asp Ala Ile Trp
 65                  70                  75                  80

att act cct gtt gtt gac agt ag gtggtgcgaa attgagtgaa aaaatgtgac      293
Ile Thr Pro Val Val Asp Ser Ser
                 85

tgacagggtg atagatac t gat ggg gga tac cat gga tac tgg gct aag gat    345
                      Asp Gly Gly Tyr His Gly Tyr Trp Ala Lys Asp
                           90                  95

ctg tat gcg gtc aac tcc aag tat ggt act gca gat gat ttg aag agt      393
Leu Tyr Ala Val Asn Ser Lys Tyr Gly Thr Ala Asp Asp Leu Lys Ser
100                 105                 110                 115

ctt gtc aaa tct gct cat gac aag gtaagcaaag ttcctctcct catgctgttc     447
Leu Val Lys Ser Ala His Asp Lys
                120

ttctgtgatc gccttggtct ttcttggtga tcccgtgctg actgttccag aac atg       503
                                                      Asn Met
                                                          125

tac gtc atg tgc gac gtc gtc gca aac cac atg ggc aaa ggc atc tca      551
Tyr Val Met Cys Asp Val Val Ala Asn His Met Gly Lys Gly Ile Ser
                130                 135                 140

gac cac aaa ccc tcc ccc ctc aac gaa caa agc tca tac cac acc ccc      599
Asp His Lys Pro Ser Pro Leu Asn Glu Gln Ser Ser Tyr His Thr Pro
            145                 150                 155

tgc gat atc gac tac agc aac cag acc agc atc gag cag tgc gaa atc      647
```

```
Cys Asp Ile Asp Tyr Ser Asn Gln Thr Ser Ile Glu Gln Cys Glu Ile
        160                 165                 170

gcc ggt ctt ccc gac ctc aac acc ggc agc gac acc gtc aag aaa gtc      695
Ala Gly Leu Pro Asp Leu Asn Thr Gly Ser Asp Thr Val Lys Lys Val
    175                 180                 185

ctc tac gac tgg atc aag tgg ctt gtc tcc gag tac agc ttc gac ggt      743
Leu Tyr Asp Trp Ile Lys Trp Leu Val Ser Glu Tyr Ser Phe Asp Gly
190                 195                 200                 205

atc cgc atc gat aca gtc aaa cat gtc gaa aag ccg ttc tgg cct ggt      791
Ile Arg Ile Asp Thr Val Lys His Val Glu Lys Pro Phe Trp Pro Gly
                210                 215                 220

ttt caa gac gct gct ggt gtt tac gcc atc ggt gag gtt tgg gat gga      839
Phe Gln Asp Ala Ala Gly Val Tyr Ala Ile Gly Glu Val Trp Asp Gly
            225                 230                 235

ggg cct gat tac cta gct ggt tat gca cag gtc atg cct ggt ctc ttg      887
Gly Pro Asp Tyr Leu Ala Gly Tyr Ala Gln Val Met Pro Gly Leu Leu
        240                 245                 250

aac tat gct atg tat tac ccc atg aac cgg ttc tac cag caa aaa ggc      935
Asn Tyr Ala Met Tyr Tyr Pro Met Asn Arg Phe Tyr Gln Gln Lys Gly
    255                 260                 265

gat cct tca gat gtt gtc gct atg cac gat gag att agc aac aaa ttc      983
Asp Pro Ser Asp Val Val Ala Met His Asp Glu Ile Ser Asn Lys Phe
270                 275                 280                 285

ccc gat ccc act atc ctc gga aca ttc atc gac aac cac gat aac cct     1031
Pro Asp Pro Thr Ile Leu Gly Thr Phe Ile Asp Asn His Asp Asn Pro
                290                 295                 300

cgc tgg ttg agc caa aag aac gac aaa gct ctc ctc aag aac gcc ctc     1079
Arg Trp Leu Ser Gln Lys Asn Asp Lys Ala Leu Leu Lys Asn Ala Leu
            305                 310                 315

gca tac gtc atc cta gcc cga ggt att ccc atc gtc tac tac gga aca     1127
Ala Tyr Val Ile Leu Ala Arg Gly Ile Pro Ile Val Tyr Tyr Gly Thr
        320                 325                 330

gag caa ggt tac gct ggc ggc aat gat ccc gcc aac cgt gag gat ctc     1175
Glu Gln Gly Tyr Ala Gly Gly Asn Asp Pro Ala Asn Arg Glu Asp Leu
    335                 340                 345

tgg cga agc agc ttc agc acc gac gcg gat ctg tat caa cac att tca     1223
Trp Arg Ser Ser Phe Ser Thr Asp Ala Asp Leu Tyr Gln His Ile Ser
350                 355                 360                 365

cgc ctc tcc aag gct cgc tcg gcg gtc ggt ggt ctt ggt gga aat gat     1271
Arg Leu Ser Lys Ala Arg Ser Ala Val Gly Gly Leu Gly Gly Asn Asp
                370                 375                 380

cac aag cat ctc tac tcg cag gac agc gcg tat gct tgg agt cgt gct     1319
His Lys His Leu Tyr Ser Gln Asp Ser Ala Tyr Ala Trp Ser Arg Ala
            385                 390                 395

gat ggc gat ctc att gtt ctt act ttg aac cgc gga cag ggg tac tcg     1367
Asp Gly Asp Leu Ile Val Leu Thr Leu Asn Arg Gly Gln Gly Tyr Ser
        400                 405                 410

ggg caa tac tgc ttt aac act gga aag aac aac aag act tgg gac aga     1415
Gly Gln Tyr Cys Phe Asn Thr Gly Lys Asn Asn Lys Thr Trp Asp Arg
    415                 420                 425

gtc ttt gga agt gga act gtt acc tct gat ggc aat ggt cag gtt tgt     1463
```

```
Val Phe Gly Ser Gly Thr Val Thr Ser Asp Gly Asn Gly Gln Val Cys
430                 435                 440                 445

gtt agc tac act aac ggc gag cct gaa gtc ttg gtt gcg tct agc ta      1510
Val Ser Tyr Thr Asn Gly Glu Pro Glu Val Leu Val Ala Ser Ser
                450                 455                 460

g                                                                  1511

<210> 481
<211> 88
<212> PRT
<213> Fungi

<400> 481
Met Lys Phe Ser Leu Leu Ala Thr Ile Val Ala Ser Ile Ser Pro Leu
 1               5                  10                  15
Ala Arg Ala Ala Asp Ala Asn Ala Trp Lys Ser Arg Asn Ile Tyr Phe
            20                  25                  30
Ala Leu Thr Asp Arg Val Ala Arg Ser Ala Asp Asp Asn Gly Gly Ser
        35                  40                  45
Ala Cys Gly Asn Leu Gly Asn Tyr Cys Gly Gly Thr Phe Lys Gly Leu
    50                  55                  60
Glu Ser Lys Leu Asp Tyr Ile Lys Gly Met Gly Phe Asp Ala Ile Trp
65                  70                  75                  80
Ile Thr Pro Val Val Asp Ser Ser
                85

<210> 482
<211> 35
<212> PRT
<213> Fungi

<400> 482
Asp Gly Gly Tyr His Gly Tyr Trp Ala Lys Asp Leu Tyr Ala Val Asn
 1               5                  10                  15
Ser Lys Tyr Gly Thr Ala Asp Asp Leu Lys Ser Leu Val Lys Ser Ala
            20                  25                  30
His Asp Lys
        35

<210> 483
<211> 337
<212> PRT
<213> Fungi

<400> 483
Asn Met Tyr Val Met Cys Asp Val Val Ala Asn His Met Gly Lys Gly
 1               5                  10                  15
Ile Ser Asp His Lys Pro Ser Pro Leu Asn Glu Gln Ser Ser Tyr His
            20                  25                  30
Thr Pro Cys Asp Ile Asp Tyr Ser Asn Gln Thr Ser Ile Glu Gln Cys
        35                  40                  45
Glu Ile Ala Gly Leu Pro Asp Leu Asn Thr Gly Ser Asp Thr Val Lys
    50                  55                  60
Lys Val Leu Tyr Asp Trp Ile Lys Trp Leu Val Ser Glu Tyr Ser Phe
65                  70                  75                  80
Asp Gly Ile Arg Ile Asp Thr Val Lys His Val Glu Lys Pro Phe Trp
                85                  90                  95
Pro Gly Phe Gln Asp Ala Ala Gly Val Tyr Ala Ile Gly Glu Val Trp
            100                 105                 110
Asp Gly Gly Pro Asp Tyr Leu Ala Gly Tyr Ala Gln Val Met Pro Gly
        115                 120                 125
Leu Leu Asn Tyr Ala Met Tyr Tyr Pro Met Asn Arg Phe Tyr Gln Gln
    130                 135                 140
Lys Gly Asp Pro Ser Asp Val Val Ala Met His Asp Glu Ile Ser Asn
145                 150                 155                 160
Lys Phe Pro Asp Pro Thr Ile Leu Gly Thr Phe Ile Asp Asn His Asp
```

```
                165                 170                 175
Asn Pro Arg Trp Leu Ser Gln Lys Asn Asp Lys Ala Leu Leu Lys Asn
            180                 185                 190
Ala Leu Ala Tyr Val Ile Leu Ala Arg Gly Ile Pro Ile Val Tyr Tyr
        195                 200                 205
Gly Thr Glu Gln Gly Tyr Ala Gly Gly Asn Asp Pro Ala Asn Arg Glu
    210                 215                 220
Asp Leu Trp Arg Ser Ser Phe Ser Thr Asp Ala Asp Leu Tyr Gln His
225                 230                 235                 240
Ile Ser Arg Leu Ser Lys Ala Arg Ser Ala Val Gly Gly Leu Gly Gly
                245                 250                 255
Asn Asp His Lys His Leu Tyr Ser Gln Asp Ser Ala Tyr Ala Trp Ser
            260                 265                 270
Arg Ala Asp Gly Asp Leu Ile Val Leu Thr Leu Asn Arg Gly Gln Gly
        275                 280                 285
Tyr Ser Gly Gln Tyr Cys Phe Asn Thr Gly Lys Asn Asn Lys Thr Trp
    290                 295                 300
Asp Arg Val Phe Gly Ser Gly Thr Val Thr Ser Asp Gly Asn Gly Gln
305                 310                 315                 320
Val Cys Val Ser Tyr Thr Asn Gly Glu Pro Glu Val Leu Val Ala Ser
                325                 330                 335
Ser
```

<210> 484
<211> 1383
<212> DNA
<213> Fungi

<400> 484

```
atgaagttct ctcttcttgc gaccatcgtc gctagcatta gtccgcttgc acgcgcagca      60
gatgcaaacg cttggaagtc gcgaaatatt tacttcgcac ttactgatcg tgttgcgcgg     120
agtgctgatg ataatggcgg cagtgcatgc ggaaacctcg ggaattactg tggtggaact     180
tttaagggct tagagtcaaa gcttgattat atcaagggca tgggatttga tgctatttgg     240
attactcctg ttgttgacag tagtgatggg ggataccatg gatactgggc taaggatctg     300
tatgcggtca actccaagta tggtactgca gatgatttga agagtcttgt caaatctgct     360
catgacaaga acatgtacgt catgtgcgac gtcgtcgcaa accacatggg caaaggcatc     420
tcagaccaca aaccctcccc cctcaacgaa caaagctcat accacacccc ctgcgatatc     480
gactacagca accagaccag catcgagcag tgcgaaatcg ccggtcttcc cgacctcaac     540
accggcagcg acaccgtcaa gaaagtcctc tacgactgga tcaagtggct tgtctccgag     600
tacagcttcg acggtatccg catcgataca gtcaaacatg tcgaaaagcc gttctggcct     660
ggttttcaag acgctgctgg tgtttacgcc atcggtgagg tttgggatgg agggcctgat     720
tacctagctg gttatgcaca ggtcatgcct ggtctcttga actatgctat gtattacccc     780
atgaaccggt tctaccagca aaaaggcgat ccttcagatg ttgtcgctat gcacgatgag     840
attagcaaca aattccccga tcccactatc ctcggaacat tcatcgacaa ccacgataac     900
cctcgctggt tgagccaaaa gaacgacaaa gctctcctca agaacgccct cgcatacgtc     960
atcctagccc gaggtattcc catcgtctac tacggaacag agcaaggtta cgctggcggc    1020
aatgatcccg ccaaccgtga ggatctctgg cgaagcagct tcagcaccga cgcggatctg    1080
tatcaacaca tttcacgcct ctccaaggct cgctcggcgg tcggtggtct tggtggaaat    1140
gatcacaagc atctctactc gcaggacagc gcgtatgctt ggagtcgtgc tgatggcgat    1200
ctcattgttc ttactttgaa ccgcggacag gggtactcgg ggcaatactg ctttaacact    1260
ggaaagaaca acaagacttg ggacagagtc tttggaagtg gaactgttac ctctgatggc    1320
aatggtcagg tttgtgttag ctacactaac ggcgagcctg aagtcttggt tgcgtctagc    1380
tag                                                                  1383
```

<210> 485
<211> 460
<212> PRT
<213> Fungi

<220>
<221> SIGNAL
<222> (1)...(19)

<221> DOMAIN
<222> (32)...(371)
<223> Alpha amylase, catalytic domain

<400> 485
Met Lys Phe Ser Leu Leu Ala Thr Ile Val Ala Ser Ile Ser Pro Leu
 1   5   10   15
Ala Arg Ala Ala Asp Ala Asn Ala Trp Lys Ser Arg Asn Ile Tyr Phe
            20   25   30
Ala Leu Thr Asp Arg Val Ala Arg Ser Ala Asp Asp Asn Gly Gly Ser
        35   40   45
Ala Cys Gly Asn Leu Gly Asn Tyr Cys Gly Gly Thr Phe Lys Gly Leu
    50   55   60
Glu Ser Lys Leu Asp Tyr Ile Lys Gly Met Gly Phe Asp Ala Ile Trp
65   70   75   80
Ile Thr Pro Val Val Asp Ser Ser Asp Gly Gly Tyr His Gly Tyr Trp
                85   90   95
Ala Lys Asp Leu Tyr Ala Val Asn Ser Lys Tyr Gly Thr Ala Asp Asp
            100   105   110
Leu Lys Ser Leu Val Lys Ser Ala His Asp Lys Asn Met Tyr Val Met
        115   120   125
Cys Asp Val Val Ala Asn His Met Gly Lys Gly Ile Ser Asp His Lys
    130   135   140
Pro Ser Pro Leu Asn Glu Gln Ser Ser Tyr His Thr Pro Cys Asp Ile
145   150   155   160
Asp Tyr Ser Asn Gln Thr Ser Ile Glu Gln Cys Glu Ile Ala Gly Leu
                165   170   175
Pro Asp Leu Asn Thr Gly Ser Asp Thr Val Lys Lys Val Leu Tyr Asp
            180   185   190
Trp Ile Lys Trp Leu Val Ser Glu Tyr Ser Phe Asp Gly Ile Arg Ile
        195   200   205
Asp Thr Val Lys His Val Glu Lys Pro Phe Trp Pro Gly Phe Gln Asp
    210   215   220
Ala Ala Gly Val Tyr Ala Ile Gly Glu Val Trp Asp Gly Gly Pro Asp
225   230   235   240
Tyr Leu Ala Gly Tyr Ala Gln Val Met Pro Gly Leu Leu Asn Tyr Ala
                245   250   255
Met Tyr Tyr Pro Met Asn Arg Phe Tyr Gln Gln Lys Gly Asp Pro Ser
            260   265   270
Asp Val Val Ala Met His Asp Glu Ile Ser Asn Lys Phe Pro Asp Pro
        275   280   285
Thr Ile Leu Gly Thr Phe Ile Asp Asn His Asp Asn Pro Arg Trp Leu
    290   295   300
Ser Gln Lys Asn Asp Lys Ala Leu Leu Lys Asn Ala Leu Ala Tyr Val
305   310   315   320
Ile Leu Ala Arg Gly Ile Pro Ile Val Tyr Tyr Gly Thr Glu Gln Gly
                325   330   335
Tyr Ala Gly Gly Asn Asp Pro Ala Asn Arg Glu Asp Leu Trp Arg Ser
            340   345   350
Ser Phe Ser Thr Asp Ala Asp Leu Tyr Gln His Ile Ser Arg Leu Ser
        355   360   365
Lys Ala Arg Ser Ala Val Gly Gly Leu Gly Gly Asn Asp His Lys His
    370   375   380
Leu Tyr Ser Gln Asp Ser Ala Tyr Ala Trp Ser Arg Ala Asp Gly Asp
385   390   395   400
Leu Ile Val Leu Thr Leu Asn Arg Gly Gln Gly Tyr Ser Gly Gln Tyr
                405   410   415
Cys Phe Asn Thr Gly Lys Asn Asn Lys Thr Trp Asp Arg Val Phe Gly
            420   425   430
Ser Gly Thr Val Thr Ser Asp Gly Asn Gly Gln Val Cys Val Ser Tyr
        435   440   445
Thr Asn Gly Glu Pro Glu Val Leu Val Ala Ser Ser
    450   455   460

<210> 486
<211> 2131
<212> DNA
<213> Fungi

<220>
<221> CDS
<222> (1)...(61)

```
<223> Exon

<221> CDS
<222> (108)...(356)
<223> Exon

<221> CDS
<222> (410)...(625)
<223> Exon

<221> CDS
<222> (673)...(1270)
<223> Exon

<221> CDS
<222> (1324)...(2131)
<223> Exon

<400> 486
atg aga cga aaa agt aca gac aag tac aag aaa gta tca ata aga gct       48
Met Arg Arg Lys Ser Thr Asp Lys Tyr Lys Lys Val Ser Ile Arg Ala
 1               5                  10                  15

cat ctt gct gca t gttattttgt ttgtacgaca gttgtttctt cttttcttga        101
His Leu Ala Ala
            20

cttgag gt gaa caa ttg gca ata tcc aag atg ctg ttt tct cga act gcc    151
      Cys Glu Gln Leu Ala Ile Ser Lys Met Leu Phe Ser Arg Thr Ala
                    25                  30                  35

acg att ttg tct ctg cta tgt gta caa gca act gca att tca ccg aga      199
Thr Ile Leu Ser Leu Leu Cys Val Gln Ala Thr Ala Ile Ser Pro Arg
                 40                  45                  50

ggt tcg gcc gca agt tat agt agt ctt gat tcc ttc att gct tcc gag      247
Gly Ser Ala Ala Ser Tyr Ser Ser Leu Asp Ser Phe Ile Ala Ser Glu
             55                  60                  65

aga aag att gct ctg cag ggt gtt ctc aac aat att gga cca gac gga      295
Arg Lys Ile Ala Leu Gln Gly Val Leu Asn Asn Ile Gly Pro Asp Gly
         70                  75                  80

tcg aaa gca tca gga gca ggc aat tat gtc att gca agt cct tcc aag      343
Ser Lys Ala Ser Gly Ala Gly Asn Tyr Val Ile Ala Ser Pro Ser Lys
     85                  90                  95

acc aat cct gat t gtaagaaata ccaatgttat gtttccaggc cgaatcttgc        396
Thr Asn Pro Asp
100

taattgattc cag ac ttc tat acc tgg act cga gat tct gcg tta acg tta    447
              Tyr Phe Tyr Thr Trp Thr Arg Asp Ser Ala Leu Thr Leu
                  105                 110                 115

aaa atg ctt gtt gat gaa ttt att ttt ggc aat cgt gca ctc cag cct      495
Lys Met Leu Val Asp Glu Phe Ile Phe Gly Asn Arg Ala Leu Gln Pro
            120                 125                 130

cag att gag aaa tac gtg aaa gct caa gcc atc ctt caa act gtt acg      543
Gln Ile Glu Lys Tyr Val Lys Ala Gln Ala Ile Leu Gln Thr Val Thr
        135                 140                 145

aat cct tct ggc act ttc ctt cct aac ggt tta gga ctt ggt gaa ccc      591
Asn Pro Ser Gly Thr Phe Leu Pro Asn Gly Leu Gly Leu Gly Glu Pro
    150                 155                 160

aaa tac gag gtt gac ggt act cga ttt aat ggt g gtatgtaatt             635
```

```
Lys Tyr Glu Val Asp Gly Thr Arg Phe Asn Gly
165                 170                 175

ttcagaggta acttctggca ctgctgacgt tccccag ca tgg ggt aga cct cag      689
                                        Ala Trp Gly Arg Pro Gln
                                                            180

cgc gat ggt cca gct ctt cgt gct atc gca tta atg gcc tac agc gat      737
Arg Asp Gly Pro Ala Leu Arg Ala Ile Ala Leu Met Ala Tyr Ser Asp
            185                 190                 195

tgg tta att aat cat aat cag ctt caa aga gcc aag cta gtt ata tgg      785
Trp Leu Ile Asn His Asn Gln Leu Gln Arg Ala Lys Leu Val Ile Trp
        200                 205                 210

cct ata atc tca aat gac tta tca tac att ggc gaa tat tgg aac caa      833
Pro Ile Ile Ser Asn Asp Leu Ser Tyr Ile Gly Glu Tyr Trp Asn Gln
    215                 220                 225

act gga tat gat tta tgg gaa gaa gtt cta gga tct agc ttt ttt acc      881
Thr Gly Tyr Asp Leu Trp Glu Glu Val Leu Gly Ser Ser Phe Phe Thr
230                 235                 240                 245

gta caa aat caa cac cgt gct ttt gtt gaa gcc gga aat ctg gcc aat      929
Val Gln Asn Gln His Arg Ala Phe Val Glu Ala Gly Asn Leu Ala Asn
                250                 255                 260

aga ctc ggt gtt gaa tct cca gga tat ctt aat gca tta gaa gct cta      977
Arg Leu Gly Val Glu Ser Pro Gly Tyr Leu Asn Ala Leu Glu Ala Leu
            265                 270                 275

tgc ttc ctt aat caa agt ttt tgg aac gga gaa tat atc gtt tcg aac     1025
Cys Phe Leu Asn Gln Ser Phe Trp Asn Gly Glu Tyr Ile Val Ser Asn
        280                 285                 290

atc aat gtc gat aac gga aga act ggg tta gac gga gcc tca att ctt     1073
Ile Asn Val Asp Asn Gly Arg Thr Gly Leu Asp Gly Ala Ser Ile Leu
    295                 300                 305

gga tct att tct atc ttc gat att gat ggc tct tgc gat agc cca aat     1121
Gly Ser Ile Ser Ile Phe Asp Ile Asp Gly Ser Cys Asp Ser Pro Asn
310                 315                 320                 325

tta caa ccg tgt agt agc aga gcg ttg tct aat ttt agg gca gtc gtg     1169
Leu Gln Pro Cys Ser Ser Arg Ala Leu Ser Asn Phe Arg Ala Val Val
                330                 335                 340

aac agc ttc cga ttt tat agt atc aat gct ggt att aaa aac aat tca     1217
Asn Ser Phe Arg Phe Tyr Ser Ile Asn Ala Gly Ile Lys Asn Asn Ser
            345                 350                 355

ggt atc gcc gtt gga aga tat gcc gaa gac att tat caa gga gga aat     1265
Gly Ile Ala Val Gly Arg Tyr Ala Glu Asp Ile Tyr Gln Gly Gly Asn
        360                 365                 370

cca tg gtactccatt ccacccctct gcggctcatt acctttgagg ctaactctcg       1320
Pro Trp
    375

cag g tac tta ctt acc act gcg gcc gca gaa ttt ctt tat gat gct gtt   1369
      Tyr Leu Leu Thr Thr Ala Ala Ala Glu Phe Leu Tyr Asp Ala Val
                    380                 385                 390

gca caa tgg aga aat cag aaa aaa ctt acc gtt gat tca aca tcc att     1417
Ala Gln Trp Arg Asn Gln Lys Lys Leu Thr Val Asp Ser Thr Ser Ile
                395                 400                 405

aat ttc ttc cgc gat ctt tac cca tct atc aaa ctt aag agt tac aca     1465
```

```
Asn Phe Phe Arg Asp Leu Tyr Pro Ser Ile Lys Leu Lys Ser Tyr Thr
            410                 415                 420

aac agg gac aac gaa ttt aac aag att ctt aat gtc gta aca aaa tac     1513
Asn Arg Asp Asn Glu Phe Asn Lys Ile Leu Asn Val Val Thr Lys Tyr
        425                 430                 435

gcc gac tct ttt gtc gca att gct caa aaa tat aca ccc agt agt gga     1561
Ala Asp Ser Phe Val Ala Ile Ala Gln Lys Tyr Thr Pro Ser Ser Gly
    440                 445                 450

tct tta gcc gaa caa ttc gat cgc aat aca gga gcc ccc att tct gcc     1609
Ser Leu Ala Glu Gln Phe Asp Arg Asn Thr Gly Ala Pro Ile Ser Ala
455                 460                 465                 470

aat gac ctt aca tgg agt tat gcg gca ttt gtc acc atg gct caa aga     1657
Asn Asp Leu Thr Trp Ser Tyr Ala Ala Phe Val Thr Met Ala Gln Arg
                475                 480                 485

cgt tca ggt caa tat cct gct agt tgg gga agt gcg gca gca tcg gtt     1705
Arg Ser Gly Gln Tyr Pro Ala Ser Trp Gly Ser Ala Ala Ala Ser Val
            490                 495                 500

cca ccc cca aat tgc gca tct tca tct tct caa gga gtg tat gtt cca     1753
Pro Pro Pro Asn Cys Ala Ser Ser Ser Ser Gln Gly Val Tyr Val Pro
        505                 510                 515

gcc att gca gca ggg gcc cca aat gta acg gca aca tgt caa tat cca     1801
Ala Ile Ala Ala Gly Ala Pro Asn Val Thr Ala Thr Cys Gln Tyr Pro
    520                 525                 530

gtg cga ttt tat gtt aat gcc aca aca tat tat ggt gaa aac ctc tac     1849
Val Arg Phe Tyr Val Asn Ala Thr Thr Tyr Tyr Gly Glu Asn Leu Tyr
535                 540                 545                 550

att att gga aat aca acg gat ttg gga gca tgg aac ttg aac agt gca     1897
Ile Ile Gly Asn Thr Thr Asp Leu Gly Ala Trp Asn Leu Asn Ser Ala
                555                 560                 565

tta cct atg aat gca ggg atg tac acg acg gaa aac cct gtt tgg tat     1945
Leu Pro Met Asn Ala Gly Met Tyr Thr Thr Glu Asn Pro Val Trp Tyr
            570                 575                 580

gtg gat gcg cag tta acg gct ggg gag cca gtc agt tat gtt tat gtg     1993
Val Asp Ala Gln Leu Thr Ala Gly Glu Pro Val Ser Tyr Val Tyr Val
        585                 590                 595

aga gaa caa gat tgc ggt caa gct cca att tac gag acg aat aat cga     2041
Arg Glu Gln Asp Cys Gly Gln Ala Pro Ile Tyr Glu Thr Asn Asn Arg
    600                 605                 610

aca agt gtg gtt ccg gaa tgt ggt aca ggt ggg tcg acg atc agg cgt     2089
Thr Ser Val Val Pro Glu Cys Gly Thr Gly Gly Ser Thr Ile Arg Arg
615                 620                 625                 630

gat gta tgg gtc ggt agt gga gga agt agt gga ggg tgt taa             2131
Asp Val Trp Val Gly Ser Gly Gly Ser Ser Gly Gly Cys  *
                635                 640
```

<210> 487
<211> 20
<212> PRT
<213> Fungi

<400> 487

```
Met Arg Arg Lys Ser Thr Asp Lys Tyr Lys Lys Val Ser Ile Arg Ala
 1               5                  10                  15
```

His Leu Ala Ala
20

<210> 488
<211> 83
<212> PRT
<213> Fungi

<400> 488
Cys Glu Gln Leu Ala Ile Ser Lys Met Leu Phe Ser Arg Thr Ala Thr
1 5 10 15
Ile Leu Ser Leu Leu Cys Val Gln Ala Thr Ala Ile Ser Pro Arg Gly
20 25 30
Ser Ala Ala Ser Tyr Ser Ser Leu Asp Ser Phe Ile Ala Ser Glu Arg
35 40 45
Lys Ile Ala Leu Gln Gly Val Leu Asn Asn Ile Gly Pro Asp Gly Ser
50 55 60
Lys Ala Ser Gly Ala Gly Asn Tyr Val Ile Ala Ser Pro Ser Lys Thr
65 70 75 80
Asn Pro Asp

<210> 489
<211> 72
<212> PRT
<213> Fungi

<400> 489
Tyr Phe Tyr Thr Trp Thr Arg Asp Ser Ala Leu Thr Leu Lys Met Leu
1 5 10 15
Val Asp Glu Phe Ile Phe Gly Asn Arg Ala Leu Gln Pro Gln Ile Glu
20 25 30
Lys Tyr Val Lys Ala Gln Ala Ile Leu Gln Thr Val Thr Asn Pro Ser
35 40 45
Gly Thr Phe Leu Pro Asn Gly Leu Gly Leu Gly Glu Pro Lys Tyr Glu
50 55 60
Val Asp Gly Thr Arg Phe Asn Gly
65 70

<210> 490
<211> 200
<212> PRT
<213> Fungi

<400> 490
Ala Trp Gly Arg Pro Gln Arg Asp Gly Pro Ala Leu Arg Ala Ile Ala
1 5 10 15
Leu Met Ala Tyr Ser Asp Trp Leu Ile Asn His Asn Gln Leu Gln Arg
20 25 30
Ala Lys Leu Val Ile Trp Pro Ile Ile Ser Asn Asp Leu Ser Tyr Ile
35 40 45
Gly Glu Tyr Trp Asn Gln Thr Gly Tyr Asp Leu Trp Glu Glu Val Leu
50 55 60
Gly Ser Ser Phe Phe Thr Val Gln Asn Gln His Arg Ala Phe Val Glu
65 70 75 80
Ala Gly Asn Leu Ala Asn Arg Leu Gly Val Glu Ser Pro Gly Tyr Leu
85 90 95
Asn Ala Leu Glu Ala Leu Cys Phe Leu Asn Gln Ser Phe Trp Asn Gly
100 105 110
Glu Tyr Ile Val Ser Asn Ile Asn Val Asp Asn Gly Arg Thr Gly Leu
115 120 125
Asp Gly Ala Ser Ile Leu Gly Ser Ile Ser Ile Phe Asp Ile Asp Gly
130 135 140
Ser Cys Asp Ser Pro Asn Leu Gln Pro Cys Ser Ser Arg Ala Leu Ser
145 150 155 160
Asn Phe Arg Ala Val Val Asn Ser Phe Arg Phe Tyr Ser Ile Asn Ala
165 170 175
Gly Ile Lys Asn Asn Ser Gly Ile Ala Val Gly Arg Tyr Ala Glu Asp

180 185 190
Ile Tyr Gln Gly Gly Asn Pro Trp
195 200

<210> 491
<211> 268
<212> PRT
<213> Fungi

<400> 491
Tyr Leu Leu Thr Thr Ala Ala Ala Glu Phe Leu Tyr Asp Ala Val Ala
1 5 10 15
Gln Trp Arg Asn Gln Lys Lys Leu Thr Val Asp Ser Thr Ser Ile Asn
20 25 30
Phe Phe Arg Asp Leu Tyr Pro Ser Ile Lys Leu Lys Ser Tyr Thr Asn
35 40 45
Arg Asp Asn Glu Phe Asn Lys Ile Leu Asn Val Val Thr Lys Tyr Ala
50 55 60
Asp Ser Phe Val Ala Ile Ala Gln Lys Tyr Thr Pro Ser Ser Gly Ser
65 70 75 80
Leu Ala Glu Gln Phe Asp Arg Asn Thr Gly Ala Pro Ile Ser Ala Asn
85 90 95
Asp Leu Thr Trp Ser Tyr Ala Ala Phe Val Thr Met Ala Gln Arg Arg
100 105 110
Ser Gly Gln Tyr Pro Ala Ser Trp Gly Ser Ala Ala Ala Ser Val Pro
115 120 125
Pro Pro Asn Cys Ala Ser Ser Ser Ser Gln Gly Val Tyr Val Pro Ala
130 135 140
Ile Ala Ala Gly Ala Pro Asn Val Thr Ala Thr Cys Gln Tyr Pro Val
145 150 155 160
Arg Phe Tyr Val Asn Ala Thr Thr Tyr Tyr Gly Glu Asn Leu Tyr Ile
165 170 175
Ile Gly Asn Thr Thr Asp Leu Gly Ala Trp Asn Leu Asn Ser Ala Leu
180 185 190
Pro Met Asn Ala Gly Met Tyr Thr Thr Glu Asn Pro Val Trp Tyr Val
195 200 205
Asp Ala Gln Leu Thr Ala Gly Glu Pro Val Ser Tyr Val Tyr Val Arg
210 215 220
Glu Gln Asp Cys Gly Gln Ala Pro Ile Tyr Glu Thr Asn Asn Arg Thr
225 230 235 240
Ser Val Val Pro Glu Cys Gly Thr Gly Gly Ser Thr Ile Arg Arg Asp
245 250 255
Val Trp Val Gly Ser Gly Gly Ser Ser Gly Gly Cys
260 265

<210> 492
<211> 1932
<212> DNA
<213> Fungi

<400> 492
atgagacgaa aaagtacaga caagtacaag aaagtatcaa taagagctca tcttgctgca 60
tgtgaacaat tggcaatatc caagatgctg ttttctcgaa ctgccacgat tttgtctctg 120
ctatgtgtac aagcaactgc aatttcaccg agaggttcgg ccgcaagtta tagtagtctt 180
gattccttca ttgcttccga gagaaagatt gctctgcagg gtgttctcaa caatattgga 240
ccagacggat cgaaagcatc aggagcaggc aattatgtca ttgcaagtcc ttccaagacc 300
aatcctgatt acttctatac ctggactcga gattctgcgt taacgttaaa aatgcttgtt 360
gatgaattta tttttggcaa tcgtgcactc cagcctcaga ttgagaaata cgtgaaagct 420
caagccatcc ttcaaactgt tacgaatcct tctggcactt tccttcctaa cggtttagga 480
cttggtgaac ccaaatacga ggttgacggt actcgattta atggtgcatg gggtagacct 540
cagcgcgatg gtccagctct tcgtgctatc gcattaatgg cctacagcga ttggttaatt 600
aatcataatc agcttcaaag agccaagcta gttatatggc ctataatctc aaatgactta 660
tcatacattg gcgaatattg gaaccaaact ggatatgatt tatgggaaga agttctagga 720
tctagctttt ttaccgtaca aaatcaacac cgtgcttttg ttgaagccgg aaatctggcc 780
aatagactcg gtgttgaatc tccaggatat cttaatgcat tagaagctct atgcttcctt 840
aatcaaagtt tttggaacgg agaatatatc gtttcgaaca tcaatgtcga taacggaaga 900
actgggttag acggagcctc aattcttgga tctatttcta tcttcgatat tgatggctct 960
tgcgatagcc caaatttaca accgtgtagt agcagagcgt tgtctaattt tagggcagtc 1020

gtgaacagct tccgatttta tagtatcaat gctggtatta aaaacaattc aggtatcgcc 1080
gttggaagat atgccgaaga catttatcaa ggaggaaatc catggtactt acttaccact 1140
gcggccgcag aatttcttta tgatgctgtt gcacaatgga gaaatcagaa aaaacttacc 1200
gttgattcaa catccattaa tttcttccgc gatctttacc catctatcaa acttaagagt 1260
tacacaaaca gggacaacga atttaacaag attcttaatg tcgtaacaaa atacgccgac 1320
tcttttgtcg caattgctca aaaatataca cccagtagtg gatctttagc cgaacaattc 1380
gatcgcaata caggagcccc catttctgcc aatgacctta catggagtta tgcggcattt 1440
gtcaccatgg ctcaaagacg ttcaggtcaa tatcctgcta gttggggaag tgcggcagca 1500
tcggttccac ccccaaattg cgcatcttca tcttctcaag gagtgtatgt tccagccatt 1560
gcagcagggg ccccaaatgt aacggcaaca tgtcaatatc cagtgcgatt ttatgttaat 1620
gccacaacat attatggtga aaacctctac attattggaa atacaacgga tttgggagca 1680
tggaacttga acagtgcatt acctatgaat gcagggatgt acacgacgga aaaccctgtt 1740
tggtatgtgg atgcgcagtt aacggctggg gagccagtca gttatgttta tgtgagagaa 1800
caagattgcg gtcaagctcc aatttacgag acgaataatc gaacaagtgt ggttccggaa 1860
tgtggtacag gtgggtcgac gatcaggcgt gatgtatggg tcggtagtgg aggaagtagt 1920
ggagggtgtt aa 1932

<210> 493
<211> 643
<212> PRT
<213> Fungi

<220>
<221> DOMAIN
<222> (532)...(628)
<223> Starch binding domain

<221> DOMAIN
<222> (69)...(485)
<223> Glycosyl hydrolases family 15

<221> SIGNAL
<222> (1)...(54)

<400> 493
Met Arg Arg Lys Ser Thr Asp Lys Tyr Lys Lys Val Ser Ile Arg Ala
 1   5   10   15
His Leu Ala Ala Cys Glu Gln Leu Ala Ile Ser Lys Met Leu Phe Ser
            20   25   30
Arg Thr Ala Thr Ile Leu Ser Leu Leu Cys Val Gln Ala Thr Ala Ile
        35   40   45
Ser Pro Arg Gly Ser Ala Ala Ser Tyr Ser Ser Leu Asp Ser Phe Ile
    50   55   60
Ala Ser Glu Arg Lys Ile Ala Leu Gln Gly Val Leu Asn Asn Ile Gly
65   70   75   80
Pro Asp Gly Ser Lys Ala Ser Gly Ala Gly Asn Tyr Val Ile Ala Ser
                85   90   95
Pro Ser Lys Thr Asn Pro Asp Tyr Phe Tyr Thr Trp Thr Arg Asp Ser
            100   105   110
Ala Leu Thr Leu Lys Met Leu Val Asp Glu Phe Ile Phe Gly Asn Arg
        115   120   125
Ala Leu Gln Pro Gln Ile Glu Lys Tyr Val Lys Ala Gln Ala Ile Leu
    130   135   140
Gln Thr Val Thr Asn Pro Ser Gly Thr Phe Leu Pro Asn Gly Leu Gly
145   150   155   160
Leu Gly Glu Pro Lys Tyr Glu Val Asp Gly Thr Arg Phe Asn Gly Ala
                165   170   175
Trp Gly Arg Pro Gln Arg Asp Gly Pro Ala Leu Arg Ala Ile Ala Leu
            180   185   190
Met Ala Tyr Ser Asp Trp Leu Ile Asn His Asn Gln Leu Gln Arg Ala
        195   200   205
Lys Leu Val Ile Trp Pro Ile Ile Ser Asn Asp Leu Ser Tyr Ile Gly
    210   215   220
Glu Tyr Trp Asn Gln Thr Gly Tyr Asp Leu Trp Glu Glu Val Leu Gly
225   230   235   240
Ser Ser Phe Phe Thr Val Gln Asn Gln His Arg Ala Phe Val Glu Ala
                245   250   255
Gly Asn Leu Ala Asn Arg Leu Gly Val Glu Ser Pro Gly Tyr Leu Asn

```
            260                 265                 270
Ala Leu Glu Ala Leu Cys Phe Leu Asn Gln Ser Phe Trp Asn Gly Glu
        275                 280                 285
Tyr Ile Val Ser Asn Ile Asn Val Asp Asn Gly Arg Thr Gly Leu Asp
    290                 295                 300
Gly Ala Ser Ile Leu Gly Ser Ile Ser Ile Phe Asp Ile Asp Gly Ser
305                 310                 315                 320
Cys Asp Ser Pro Asn Leu Gln Pro Cys Ser Ser Arg Ala Leu Ser Asn
                325                 330                 335
Phe Arg Ala Val Val Asn Ser Phe Arg Phe Tyr Ser Ile Asn Ala Gly
            340                 345                 350
Ile Lys Asn Asn Ser Gly Ile Ala Val Gly Arg Tyr Ala Glu Asp Ile
        355                 360                 365
Tyr Gln Gly Gly Asn Pro Trp Tyr Leu Leu Thr Thr Ala Ala Ala Glu
    370                 375                 380
Phe Leu Tyr Asp Ala Val Ala Gln Trp Arg Asn Gln Lys Lys Leu Thr
385                 390                 395                 400
Val Asp Ser Thr Ser Ile Asn Phe Phe Arg Asp Leu Tyr Pro Ser Ile
                405                 410                 415
Lys Leu Lys Ser Tyr Thr Asn Arg Asp Asn Glu Phe Asn Lys Ile Leu
            420                 425                 430
Asn Val Val Thr Lys Tyr Ala Asp Ser Phe Val Ala Ile Ala Gln Lys
        435                 440                 445
Tyr Thr Pro Ser Ser Gly Ser Leu Ala Glu Gln Phe Asp Arg Asn Thr
    450                 455                 460
Gly Ala Pro Ile Ser Ala Asn Asp Leu Thr Trp Ser Tyr Ala Ala Phe
465                 470                 475                 480
Val Thr Met Ala Gln Arg Arg Ser Gly Gln Tyr Pro Ala Ser Trp Gly
                485                 490                 495
Ser Ala Ala Ala Ser Val Pro Pro Pro Asn Cys Ala Ser Ser Ser Ser
            500                 505                 510
Gln Gly Val Tyr Val Pro Ala Ile Ala Ala Gly Ala Pro Asn Val Thr
        515                 520                 525
Ala Thr Cys Gln Tyr Pro Val Arg Phe Tyr Val Asn Ala Thr Thr Tyr
    530                 535                 540
Tyr Gly Glu Asn Leu Tyr Ile Ile Gly Asn Thr Thr Asp Leu Gly Ala
545                 550                 555                 560
Trp Asn Leu Asn Ser Ala Leu Pro Met Asn Ala Gly Met Tyr Thr Thr
                565                 570                 575
Glu Asn Pro Val Trp Tyr Val Asp Ala Gln Leu Thr Ala Gly Glu Pro
            580                 585                 590
Val Ser Tyr Val Tyr Val Arg Glu Gln Asp Cys Gly Gln Ala Pro Ile
        595                 600                 605
Tyr Glu Thr Asn Asn Arg Thr Ser Val Val Pro Glu Cys Gly Thr Gly
    610                 615                 620
Gly Ser Thr Ile Arg Arg Asp Val Trp Val Gly Ser Gly Gly Ser Ser
625                 630                 635                 640
Gly Gly Cys
```

<210> 494
<211> 1586
<212> DNA
<213> Fungi

<220>
<221> CDS
<222> (1)...(225)
<223> Exon

<221> CDS
<222> (283)...(450)
<223> Exon

<221> CDS
<222> (501)...(1586)
<223> Exon

<400> 494

```
atg gga ttt tcg aaa atg ttg ttg ggc gcc ttg ata ggc atc gct tcg       48
Met Gly Phe Ser Lys Met Leu Leu Gly Ala Leu Ile Gly Ile Ala Ser
 1               5                  10                  15

tta aac ggc gtt caa tca gca acg ccc gac gaa tgg agg act cgt tca       96
Leu Asn Gly Val Gln Ser Ala Thr Pro Asp Glu Trp Arg Thr Arg Ser
            20                  25                  30

ata tat caa gta ttt aca gat cga ttt gca aga aca gat gtt tca aat      144
Ile Tyr Gln Val Phe Thr Asp Arg Phe Ala Arg Thr Asp Val Ser Asn
         35                  40                  45

act acg gat tgc cct tct caa act cgg gga tac tgt ggt ggg aca tgg      192
Thr Thr Asp Cys Pro Ser Gln Thr Arg Gly Tyr Cys Gly Gly Thr Trp
     50                  55                  60

caa ggc ata att aac aag ctc gat tat atc cag gtatgagagt agtgcaatca    245
Gln Gly Ile Ile Asn Lys Leu Asp Tyr Ile Gln
 65                  70                  75

catattggta aaattacaag tgctaatagg ctattag gat atg ggg ttc aca gca     300
                                        Asp Met Gly Phe Thr Ala
                                                             80

att tgg ata tca cca gtt gtt gaa caa gtt ggg gat cca ggt cgc ggc      348
Ile Trp Ile Ser Pro Val Val Glu Gln Val Gly Asp Pro Gly Arg Gly
            85                  90                  95

ttt cat ggg tat tcg gct cag aat ttg tat ggt cta aac agt cac ttc      396
Phe His Gly Tyr Ser Ala Gln Asn Leu Tyr Gly Leu Asn Ser His Phe
        100                 105                 110

gga gac gca gcg gat ttg aaa gcg cta gct acc gcg ctc cat gat cga      444
Gly Asp Ala Ala Asp Leu Lys Ala Leu Ala Thr Ala Leu His Asp Arg
    115                 120                 125

gga atg gtgagactga acaagaaatt tgcacagata aagaaagcta attctggtag       500
Gly Met
130

tat ttg atg gtt gat gtc gtg gcg aac cat atg gga tct gat gat act      548
Tyr Leu Met Val Asp Val Val Ala Asn His Met Gly Ser Asp Asp Thr
            135                 140                 145

gcg cat aac gtt gat ttc agc atc atg aat cca ttc aat gac agc aaa      596
Ala His Asn Val Asp Phe Ser Ile Met Asn Pro Phe Asn Asp Ser Lys
        150                 155                 160

tac ttt cac agt gtc tgc ttt atc aat gat tac aac aat cag aca aat      644
Tyr Phe His Ser Val Cys Phe Ile Asn Asp Tyr Asn Asn Gln Thr Asn
    165                 170                 175

gtt gaa cta tgc gag ctt gga acg gag cgg tat cca tta cct gat cta      692
Val Glu Leu Cys Glu Leu Gly Thr Glu Arg Tyr Pro Leu Pro Asp Leu
180                 185                 190                 195

aac act act cga caa gat gtt cga gat ctc cat act act tgg atc aag      740
Asn Thr Thr Arg Gln Asp Val Arg Asp Leu His Thr Thr Trp Ile Lys
                200                 205                 210

tca cta gta gcc aac tac tca att gac gga ttg aga gtg gac act gtt      788
Ser Leu Val Ala Asn Tyr Ser Ile Asp Gly Leu Arg Val Asp Thr Val
            215                 220                 225

aga cat gta gag aag gat ttt tgg cct tta ttc aat gaa acc gcc ggt      836
Arg His Val Glu Lys Asp Phe Trp Pro Leu Phe Asn Glu Thr Ala Gly
        230                 235                 240
```

```
gtt tat tgt gtt ggc gaa gtg gca gat gga gac gta gat tat cta tgt      884
Val Tyr Cys Val Gly Glu Val Ala Asp Gly Asp Val Asp Tyr Leu Cys
    245                 250                 255

ccc tac caa gat tac ata gat ggg ctt ctc tcc tat gcc tca tac ttc      932
Pro Tyr Gln Asp Tyr Ile Asp Gly Leu Leu Ser Tyr Ala Ser Tyr Phe
260                 265                 270                 275

caa ctc aca aaa ttc ttt tcc aac acc tct gca aca tca gag aac tta      980
Gln Leu Thr Lys Phe Phe Ser Asn Thr Ser Ala Thr Ser Glu Asn Leu
                280                 285                 290

ata gga caa tta caa tcc cag aat tat caa tgt aaa gac acg aca ctt     1028
Ile Gly Gln Leu Gln Ser Gln Asn Tyr Gln Cys Lys Asp Thr Thr Leu
            295                 300                 305

ctg gga tcc ttc acc gag aac cat gat caa ccg cgc ttc gga aat tac     1076
Leu Gly Ser Phe Thr Glu Asn His Asp Gln Pro Arg Phe Gly Asn Tyr
        310                 315                 320

act agt gac ctc acg tta gct aaa aac atc atc aca tat acg atg tta     1124
Thr Ser Asp Leu Thr Leu Ala Lys Asn Ile Ile Thr Tyr Thr Met Leu
    325                 330                 335

gca gat ggt ata cca att atc tat caa ggg caa gaa cag cat ttc cat     1172
Ala Asp Gly Ile Pro Ile Ile Tyr Gln Gly Gln Glu Gln His Phe His
340                 345                 350                 355

ggc gca act gac cct tat gat cgt gaa ccc ata tgg cct aca ggt tac     1220
Gly Ala Thr Asp Pro Tyr Asp Arg Glu Pro Ile Trp Pro Thr Gly Tyr
                360                 365                 370

gat act act tcc cca ctt tat gtc ctg gtc aaa caa ttg aac gct ata     1268
Asp Thr Thr Ser Pro Leu Tyr Val Leu Val Lys Gln Leu Asn Ala Ile
            375                 380                 385

cgt tct ttg gcg ata gca cgc tcg gat aca tat gct act tat caa acc     1316
Arg Ser Leu Ala Ile Ala Arg Ser Asp Thr Tyr Ala Thr Tyr Gln Thr
        390                 395                 400

caa atc gct tac tca gat ccg cac aat att gcg ttt cgg agg gga gat     1364
Gln Ile Ala Tyr Ser Asp Pro His Asn Ile Ala Phe Arg Arg Gly Asp
    405                 410                 415

gag aaa tgc atg agt cta atg gtg ctg aac aat atc ggt gag agt gca     1412
Glu Lys Cys Met Ser Leu Met Val Leu Asn Asn Ile Gly Glu Ser Ala
420                 425                 430                 435

gag gat tat aca gtc gag att gag aat gtg ggg ttt gaa gct gga tcg     1460
Glu Asp Tyr Thr Val Glu Ile Glu Asn Val Gly Phe Glu Ala Gly Ser
                440                 445                 450

ata gtg acg gat gta ttg agt tgt aga aat gta atg gta gat gag tat     1508
Ile Val Thr Asp Val Leu Ser Cys Arg Asn Val Met Val Asp Glu Tyr
            455                 460                 465

ggt ggg atg gaa gtg cca ttt gtg agc ggg tta cct agt gta agt cta     1556
Gly Gly Met Glu Val Pro Phe Val Ser Gly Leu Pro Ser Val Ser Leu
        470                 475                 480

atc ctc tcc atg aga ctg ttt aaa atc taa                            1586
Ile Leu Ser Met Arg Leu Phe Lys Ile  *
    485                 490
```

<210> 495
<211> 75

<212> PRT
<213> Fungi

<400> 495
Met Gly Phe Ser Lys Met Leu Leu Gly Ala Leu Ile Gly Ile Ala Ser
 1   5   10   15
Leu Asn Gly Val Gln Ser Ala Thr Pro Asp Glu Trp Arg Thr Arg Ser
            20   25   30
Ile Tyr Gln Val Phe Thr Asp Arg Phe Ala Arg Thr Asp Val Ser Asn
        35   40   45
Thr Thr Asp Cys Pro Ser Gln Thr Arg Gly Tyr Cys Gly Gly Thr Trp
    50   55   60
Gln Gly Ile Ile Asn Lys Leu Asp Tyr Ile Gln
65   70   75

<210> 496
<211> 56
<212> PRT
<213> Fungi

<400> 496
Asp Met Gly Phe Thr Ala Ile Trp Ile Ser Pro Val Val Glu Gln Val
 1   5   10   15
Gly Asp Pro Gly Arg Gly Phe His Gly Tyr Ser Ala Gln Asn Leu Tyr
            20   25   30
Gly Leu Asn Ser His Phe Gly Asp Ala Ala Asp Leu Lys Ala Leu Ala
        35   40   45
Thr Ala Leu His Asp Arg Gly Met
    50   55

<210> 497
<211> 361
<212> PRT
<213> Fungi

<400> 497
Tyr Leu Met Val Asp Val Val Ala Asn His Met Gly Ser Asp Asp Thr
 1   5   10   15
Ala His Asn Val Asp Phe Ser Ile Met Asn Pro Phe Asn Asp Ser Lys
            20   25   30
Tyr Phe His Ser Val Cys Phe Ile Asn Asp Tyr Asn Asn Gln Thr Asn
        35   40   45
Val Glu Leu Cys Glu Leu Gly Thr Glu Arg Tyr Pro Leu Pro Asp Leu
    50   55   60
Asn Thr Thr Arg Gln Asp Val Arg Asp Leu His Thr Thr Trp Ile Lys
65   70   75   80
Ser Leu Val Ala Asn Tyr Ser Ile Asp Gly Leu Arg Val Asp Thr Val
                85   90   95
Arg His Val Glu Lys Asp Phe Trp Pro Leu Phe Asn Glu Thr Ala Gly
            100   105   110
Val Tyr Cys Val Gly Glu Val Ala Asp Gly Asp Val Asp Tyr Leu Cys
        115   120   125
Pro Tyr Gln Asp Tyr Ile Asp Gly Leu Leu Ser Tyr Ala Ser Tyr Phe
    130   135   140
Gln Leu Thr Lys Phe Phe Ser Asn Thr Ser Ala Thr Ser Glu Asn Leu
145   150   155   160
Ile Gly Gln Leu Gln Ser Gln Asn Tyr Gln Cys Lys Asp Thr Thr Leu
                165   170   175
Leu Gly Ser Phe Thr Glu Asn His Asp Gln Pro Arg Phe Gly Asn Tyr
            180   185   190
Thr Ser Asp Leu Thr Leu Ala Lys Asn Ile Ile Thr Tyr Thr Met Leu
        195   200   205
Ala Asp Gly Ile Pro Ile Ile Tyr Gln Gly Gln Glu Gln His Phe His
    210   215   220
Gly Ala Thr Asp Pro Tyr Asp Arg Glu Pro Ile Trp Pro Thr Gly Tyr
225   230   235   240
Asp Thr Thr Ser Pro Leu Tyr Val Leu Val Lys Gln Leu Asn Ala Ile
                245   250   255

```
Arg Ser Leu Ala Ile Ala Arg Ser Asp Thr Tyr Ala Thr Tyr Gln Thr
            260                 265                 270
Gln Ile Ala Tyr Ser Asp Pro His Asn Ile Ala Phe Arg Arg Gly Asp
        275                 280                 285
Glu Lys Cys Met Ser Leu Met Val Leu Asn Asn Ile Gly Glu Ser Ala
    290                 295                 300
Glu Asp Tyr Thr Val Glu Ile Glu Asn Val Gly Phe Glu Ala Gly Ser
305                 310                 315                 320
Ile Val Thr Asp Val Leu Ser Cys Arg Asn Val Met Val Asp Glu Tyr
                325                 330                 335
Gly Gly Met Glu Val Pro Phe Val Ser Gly Leu Pro Ser Val Ser Leu
            340                 345                 350
Ile Leu Ser Met Arg Leu Phe Lys Ile
        355                 360
```

<210> 498
<211> 1479
<212> DNA
<213> Fungi

<400> 498

```
atgggatttt cgaaaatgtt gttgggcgcc ttgataggca tcgcttcgtt aaacggcgtt      60
caatcagcaa cgcccgacga atggaggact cgttcaatat atcaagtatt tacagatcga     120
tttgcaagaa cagatgtttc aaatactacg gattgccctt ctcaaactcg gggatactgt     180
ggtgggacat ggcaaggcat aattaacaag ctcgattata tccaggatat ggggttcaca     240
gcaatttgga tatcaccagt tgttgaacaa gttggggatc caggtcgcgg ctttcatggg     300
tattcggctc agaatttgta tggtctaaac agtcacttcg gagacgcagc ggatttgaaa     360
gcgctagcta ccgcgctcca tgatcgagga atgtatttga tggttgatgt cgtggcgaac     420
catatgggat ctgatgatac tgcgcataac gttgatttca gcatcatgaa tccattcaat     480
gacagcaaat actttcacag tgtctgcttt atcaatgatt acaacaatca gacaaatgtt     540
gaactatgcg agcttggaac ggagcggtat ccattacctg atctaaacac tactcgacaa     600
gatgttcgag atctccatac tacttggatc aagtcactag tagccaacta ctcaattgac     660
ggattgagag tggacactgt tagacatgta gagaaggatt tttggccttt attcaatgaa     720
accgccggtg tttattgtgt tggcgaagtg gcagatggag acgtagatta tctatgtccc     780
taccaagatt acatagatgg gcttctctcc tatgcctcat acttccaact cacaaaattc     840
ttttccaaca cctctgcaac atcagagaac ttaataggac aattacaatc ccagaattat     900
caatgtaaag acacgacact tctgggatcc ttcaccgaga accatgatca accgcgcttc     960
ggaaattaca ctagtgacct cacgttagct aaaaacatca tcacatatac gatgttagca    1020
gatggtatac caattatcta tcaagggcaa gaacagcatt tccatggcgc aactgaccct    1080
tatgatcgtg aacccatatg gcctacaggt tacgatacta cttccccact ttatgtcctg    1140
gtcaaacaat tgaacgctat acgttctttg gcgatagcac gctcggatac atatgctact    1200
tatcaaaccc aaatcgctta ctcagatccg cacaatattg cgtttcggag gggagatgag    1260
aaatgcatga gtctaatggt gctgaacaat atcggtgaga gtgcagagga ttatacagtc    1320
gagattgaga atgtggggtt tgaagctgga tcgatagtga cggatgtatt gagttgtaga    1380
aatgtaatgg tagatgagta tggtgggatg gaagtgccat ttgtgagcgg gttacctagt    1440
gtaagtctaa tcctctccat gagactgttt aaaatctaa                           1479
```

<210> 499
<211> 492
<212> PRT
<213> Fungi

<220>
<221> SIGNAL
<222> (1)...(22)

<221> DOMAIN
<222> (35)...(388)
<223> Alpha amylase, catalytic domain

<400> 499

```
Met Gly Phe Ser Lys Met Leu Leu Gly Ala Leu Ile Gly Ile Ala Ser
 1               5                  10                  15
Leu Asn Gly Val Gln Ser Ala Thr Pro Asp Glu Trp Arg Thr Arg Ser
            20                  25                  30
Ile Tyr Gln Val Phe Thr Asp Arg Phe Ala Arg Thr Asp Val Ser Asn
        35                  40                  45
Thr Thr Asp Cys Pro Ser Gln Thr Arg Gly Tyr Cys Gly Gly Thr Trp
```

```
    50                  55                  60
Gln Gly Ile Ile Asn Lys Leu Asp Tyr Ile Gln Asp Met Gly Phe Thr
65                  70                  75                  80
Ala Ile Trp Ile Ser Pro Val Val Glu Gln Val Gly Asp Pro Gly Arg
                85                  90                  95
Gly Phe His Gly Tyr Ser Ala Gln Asn Leu Tyr Gly Leu Asn Ser His
            100                 105                 110
Phe Gly Asp Ala Ala Asp Leu Lys Ala Leu Ala Thr Ala Leu His Asp
        115                 120                 125
Arg Gly Met Tyr Leu Met Val Asp Val Val Ala Asn His Met Gly Ser
    130                 135                 140
Asp Asp Thr Ala His Asn Val Asp Phe Ser Ile Met Asn Pro Phe Asn
145                 150                 155                 160
Asp Ser Lys Tyr Phe His Ser Val Cys Phe Ile Asn Asp Tyr Asn Asn
                165                 170                 175
Gln Thr Asn Val Glu Leu Cys Glu Leu Gly Thr Glu Arg Tyr Pro Leu
            180                 185                 190
Pro Asp Leu Asn Thr Thr Arg Gln Asp Val Arg Asp Leu His Thr Thr
        195                 200                 205
Trp Ile Lys Ser Leu Val Ala Asn Tyr Ser Ile Asp Gly Leu Arg Val
    210                 215                 220
Asp Thr Val Arg His Val Glu Lys Asp Phe Trp Pro Leu Phe Asn Glu
225                 230                 235                 240
Thr Ala Gly Val Tyr Cys Val Gly Glu Val Ala Asp Gly Asp Val Asp
                245                 250                 255
Tyr Leu Cys Pro Tyr Gln Asp Tyr Ile Asp Gly Leu Leu Ser Tyr Ala
            260                 265                 270
Ser Tyr Phe Gln Leu Thr Lys Phe Phe Ser Asn Thr Ser Ala Thr Ser
        275                 280                 285
Glu Asn Leu Ile Gly Gln Leu Gln Ser Gln Asn Tyr Gln Cys Lys Asp
    290                 295                 300
Thr Thr Leu Leu Gly Ser Phe Thr Glu Asn His Asp Gln Pro Arg Phe
305                 310                 315                 320
Gly Asn Tyr Thr Ser Asp Leu Thr Leu Ala Lys Asn Ile Ile Thr Tyr
                325                 330                 335
Thr Met Leu Ala Asp Gly Ile Pro Ile Ile Tyr Gln Gly Gln Glu Gln
            340                 345                 350
His Phe His Gly Ala Thr Asp Pro Tyr Asp Arg Glu Pro Ile Trp Pro
        355                 360                 365
Thr Gly Tyr Asp Thr Thr Ser Pro Leu Tyr Val Leu Val Lys Gln Leu
    370                 375                 380
Asn Ala Ile Arg Ser Leu Ala Ile Ala Arg Ser Asp Thr Tyr Ala Thr
385                 390                 395                 400
Tyr Gln Thr Gln Ile Ala Tyr Ser Asp Pro His Asn Ile Ala Phe Arg
                405                 410                 415
Arg Gly Asp Glu Lys Cys Met Ser Leu Met Val Leu Asn Asn Ile Gly
            420                 425                 430
Glu Ser Ala Glu Asp Tyr Thr Val Glu Ile Glu Asn Val Gly Phe Glu
        435                 440                 445
Ala Gly Ser Ile Val Thr Asp Val Leu Ser Cys Arg Asn Val Met Val
    450                 455                 460
Asp Glu Tyr Gly Gly Met Glu Val Pro Phe Val Ser Gly Leu Pro Ser
465                 470                 475                 480
Val Ser Leu Ile Leu Ser Met Arg Leu Phe Lys Ile
                485                 490
```

<210> 500
<211> 2236
<212> DNA
<213> Fungi

<220>
<221> CDS
<222> (1)...(63)
<223> Exon

<221> CDS
<222> (114)...(207)

<223> Exon

<221> CDS
<222> (264)...(287)
<223> Exon

<221> CDS
<222> (341)...(378)
<223> Exon

<221> CDS
<222> (435)...(560)
<223> Exon

<221> CDS
<222> (623)...(1521)
<223> Exon

<221> CDS
<222> (1571)...(1973)
<223> Exon

<221> CDS
<222> (2051)...(2236)
<223> Exon

<400> 500
atg agt caa gtt tct gga aaa gtg ccc gac agg aag tgg tgg aag gaa 48
Met Ser Gln Val Ser Gly Lys Val Pro Asp Arg Lys Trp Trp Lys Glu
1 5 10 15

gct att gtc tac cag gtttgaacta cttttctctt gtcactgcat gcgacttata 103
Ala Ile Val Tyr Gln
20

cggcggctag atc tat ccc gcg tcg ttt ctt gat acc gat tct aat ggc 152
Ile Tyr Pro Ala Ser Phe Leu Asp Thr Asp Ser Asn Gly
25 30

gtc ggg aac atc aat ggg atc aca gct aag ctg gac tac ctc aag gaa 200
Val Gly Asn Ile Asn Gly Ile Thr Ala Lys Leu Asp Tyr Leu Lys Glu
35 40 45 50

ttg ggc g gtaagaactg acctccctcc tcctatgctc atgatctaac gcgtgtgtgt 257
Leu Gly

tcacag tg gac att ctc tgg att tca ccc g gtaagctcta tcccagccgt 307
Val Asp Ile Leu Trp Ile Ser Pro
55 60

agcgacgcgc ttccagtcta acacgcgctc tag tt tac gaa agc ccc cag aag 360
Val Tyr Glu Ser Pro Gln Lys
65

gac atg gga tac gac att gtacgacacc catctcttta taaaaattga 408
Asp Met Gly Tyr Asp Ile
70

tgtagcaccc actgaccttg tcgcag tcg aat tac cgc gat atc cat cgc ccc 461
Ser Asn Tyr Arg Asp Ile His Arg Pro
75 80

tac ggc acc atg caa gac gtg cac cac ttg atc aac gag ctc aaa gca 509
Tyr Gly Thr Met Gln Asp Val His His Leu Ile Asn Glu Leu Lys Ala
85 90 95

cgc gac atg aaa ctg atc atg gat ctg gtg gtc aat cac acc tcg acc 557

```
Arg Asp Met Lys Leu Ile Met Asp Leu Val Val Asn His Thr Ser Thr
    100                 105                 110

gag gtaagatagc ctttcgaact ctgtatcctg tttaacgtga ctgatttcac          610
Glu
115

ctgacgctat ag cat ccc tgg ttc atc gac tca gcc tcg tcc gca gta agc    661
             His Pro Trp Phe Ile Asp Ser Ala Ser Ser Ala Val Ser
                         120                 125

gcg cga cga gac tgg tac atc tgg cgc aag cca aag tac gac gag cag      709
Ala Arg Arg Asp Trp Tyr Ile Trp Arg Lys Pro Lys Tyr Asp Glu Gln
    130                 135                 140

ggc aat cgg cag cca ccg aat aac tgg tgc agc ctc ttt gac gag act      757
Gly Asn Arg Gln Pro Pro Asn Asn Trp Cys Ser Leu Phe Asp Glu Thr
145                 150                 155                 160

gaa tca gcc tgg acc tac gac gct aag acg gat gag tac tat ctc tcg      805
Glu Ser Ala Trp Thr Tyr Asp Ala Lys Thr Asp Glu Tyr Tyr Leu Ser
                165                 170                 175

ctg ttt agt ccg ttc cag gca gat ctt aac tgg gag acc cca ttc gtc      853
Leu Phe Ser Pro Phe Gln Ala Asp Leu Asn Trp Glu Thr Pro Phe Val
            180                 185                 190

aga gag gaa gtc tgc gac atc ctg cgc ttc tgg ctc gac aag ggc gtc      901
Arg Glu Glu Val Cys Asp Ile Leu Arg Phe Trp Leu Asp Lys Gly Val
        195                 200                 205

tct ggt ttc cgg atg gac gtc atc aac ctc atc tcc aag gac cag agg      949
Ser Gly Phe Arg Met Asp Val Ile Asn Leu Ile Ser Lys Asp Gln Arg
    210                 215                 220

ttc cct gac gca gag atc cgc cac ccg ggc agg aag tac cag ccg ggt      997
Phe Pro Asp Ala Glu Ile Arg His Pro Gly Arg Lys Tyr Gln Pro Gly
225                 230                 235                 240

gaa tgc tac ttc gcc aac ggg gtc cga ctc atg gac tat ctc cag gag     1045
Glu Cys Tyr Phe Ala Asn Gly Val Arg Leu Met Asp Tyr Leu Gln Glu
                245                 250                 255

atg aag acc gcc gta ttc tca aag tat gac gcc gtc acg gtc ggc gag     1093
Met Lys Thr Ala Val Phe Ser Lys Tyr Asp Ala Val Thr Val Gly Glu
            260                 265                 270

atg ccg tac ctt gaa gat gaa gag cag aga ctc agg atg gtc gcg gct     1141
Met Pro Tyr Leu Glu Asp Glu Glu Gln Arg Leu Arg Met Val Ala Ala
        275                 280                 285

gaa gaa ggc gtg ctg aac atg atc ttc act ttt gag atg att ggc ctc     1189
Glu Glu Gly Val Leu Asn Met Ile Phe Thr Phe Glu Met Ile Gly Leu
    290                 295                 300

gac atc gtc ccc gaa aaa ggc cgc ttc agc aat aaa gct tgg aag gtc     1237
Asp Ile Val Pro Glu Lys Gly Arg Phe Ser Asn Lys Ala Trp Lys Val
305                 310                 315                 320

tgt gag ctg aag gac ata gtt gcc aag gcc tgc aag ctc gtt aac cag     1285
Cys Glu Leu Lys Asp Ile Val Ala Lys Ala Cys Lys Leu Val Asn Gln
                325                 330                 335

gac ggg tgg cat acg ctc ttc tgc gag aat cat gac cag ccg cgt tcc     1333
Asp Gly Trp His Thr Leu Phe Cys Glu Asn His Asp Gln Pro Arg Ser
            340                 345                 350

gtc acg cgt ttc tgc gac gac tcg gat gag cat cgt gtc gca ggc acc     1381
```

```
Val Thr Arg Phe Cys Asp Asp Ser Asp Glu His Arg Val Ala Gly Thr
        355                 360                 365

aag ctc ctc tcc atc atg cag aca agc ctt cct ggt acc ctg tat gtc     1429
Lys Leu Leu Ser Ile Met Gln Thr Ser Leu Pro Gly Thr Leu Tyr Val
    370                 375                 380

tac cag ggc gaa gag ctg ggt atg cga aac gtg ccc aag tcc tgg gga     1477
Tyr Gln Gly Glu Glu Leu Gly Met Arg Asn Val Pro Lys Ser Trp Gly
385                 390                 395                 400

ttt gaa gaa tac ata gac att gag tct gtc tcc tac atc aag aa          1521
Phe Glu Glu Tyr Ile Asp Ile Glu Ser Val Ser Tyr Ile Lys Asn
                405                 410                 415

gtaagtctca attctcttat cgactttaac atctactgac atggtatag t gtc tgc     1577
                                                    Val Cys

gct caa ttc cca gaa ggg tct ccg gaa aga aaa gag gct aag aat ctt     1625
Ala Gln Phe Pro Glu Gly Ser Pro Glu Arg Lys Glu Ala Lys Asn Leu
        420                 425                 430

ctg cgc ctc aaa gct cgt gat aat gcc agg acc cca atg cag tgg gac     1673
Leu Arg Leu Lys Ala Arg Asp Asn Ala Arg Thr Pro Met Gln Trp Asp
    435                 440                 445

tct aca tct aac ggc ggc ttc tgt ccc gag gga gtc aag ccc tgg atg     1721
Ser Thr Ser Asn Gly Gly Phe Cys Pro Glu Gly Val Lys Pro Trp Met
450                 455                 460                 465

cgc gtc aac gac gac tac acg atg gtc aac gca gct cta cag act tcc     1769
Arg Val Asn Asp Asp Tyr Thr Met Val Asn Ala Ala Leu Gln Thr Ser
                470                 475                 480

gca ggc cag gca acc gac cgc acc atg ctt gtg tca ccc tat agg ttc     1817
Ala Gly Gln Ala Thr Asp Arg Thr Met Leu Val Ser Pro Tyr Arg Phe
            485                 490                 495

tgg cag cga agc atc cag ata cgc aag aag ttc aaa gac ctg ctc gtg     1865
Trp Gln Arg Ser Ile Gln Ile Arg Lys Lys Phe Lys Asp Leu Leu Val
        500                 505                 510

tac gga gac ctt gag att att gat ggc acg cac gcc aac gtc ttt gcg     1913
Tyr Gly Asp Leu Glu Ile Ile Asp Gly Thr His Ala Asn Val Phe Ala
    515                 520                 525

ttc aag agg atc agc aac ggc aga cat acc atc acg atc ctt aac ttt     1961
Phe Lys Arg Ile Ser Asn Gly Arg His Thr Ile Thr Ile Leu Asn Phe
530                 535                 540                 545

agt aag gag gag gtgagtttca agtttccgga aggtgagggg gtcaggggct         2013
Ser Lys Glu Glu

gggctctcgg gagttatgat gttttgtctt tggcgag gcc gtg gac ggg gga gat    2068
                                        Ala Val Asp Gly Gly Asp
                                        550                 555

aca gct ttt acc ttg gga ggg ctt gct tgg tat tgc ttg atg gat ctc     2116
Thr Ala Phe Thr Leu Gly Gly Leu Ala Trp Tyr Cys Leu Met Asp Leu
                560                 565                 570

tca ttt tgg aat gga agt aaa aga ttg gta tca atg gct acg ata ttg     2164
Ser Phe Trp Asn Gly Ser Lys Arg Leu Val Ser Met Ala Thr Ile Leu
            575                 580                 585

tct tta tat agc cag tac aga atg cca cgg act gtt agt aac atg caa     2212
```

Ser Leu Tyr Ser Gln Tyr Arg Met Pro Arg Thr Val Ser Asn Met Gln
590 595 600

tgt ctt gtg ata cca tcc agt taa 2236
Cys Leu Val Ile Pro Ser Ser *
605 610

<210> 501
<211> 21
<212> PRT
<213> Fungi

<400> 501
Met Ser Gln Val Ser Gly Lys Val Pro Asp Arg Lys Trp Trp Lys Glu
1 5 10 15
Ala Ile Val Tyr Gln
20

<210> 502
<211> 31
<212> PRT
<213> Fungi

<400> 502
Ile Tyr Pro Ala Ser Phe Leu Asp Thr Asp Ser Asn Gly Val Gly Asn
1 5 10 15
Ile Asn Gly Ile Thr Ala Lys Leu Asp Tyr Leu Lys Glu Leu Gly
20 25 30

<210> 503
<211> 8
<212> PRT
<213> Fungi

<400> 503
Val Asp Ile Leu Trp Ile Ser Pro
1 5

<210> 504
<211> 13
<212> PRT
<213> Fungi

<400> 504
Val Tyr Glu Ser Pro Gln Lys Asp Met Gly Tyr Asp Ile
1 5 10

<210> 505
<211> 42
<212> PRT
<213> Fungi

<400> 505
Ser Asn Tyr Arg Asp Ile His Arg Pro Tyr Gly Thr Met Gln Asp Val
1 5 10 15
His His Leu Ile Asn Glu Leu Lys Ala Arg Asp Met Lys Leu Ile Met
20 25 30
Asp Leu Val Val Asn His Thr Ser Thr Glu
35 40

<210> 506
<211> 300
<212> PRT
<213> Fungi

<400> 506
His Pro Trp Phe Ile Asp Ser Ala Ser Ser Ala Val Ser Ala Arg Arg

```
 1               5                  10                  15
Asp Trp Tyr Ile Trp Arg Lys Pro Lys Tyr Asp Glu Gln Gly Asn Arg
            20                  25                  30
Gln Pro Pro Asn Asn Trp Cys Ser Leu Phe Asp Glu Thr Glu Ser Ala
        35                  40                  45
Trp Thr Tyr Asp Ala Lys Thr Asp Glu Tyr Tyr Leu Ser Leu Phe Ser
    50                  55                  60
Pro Phe Gln Ala Asp Leu Asn Trp Glu Thr Pro Phe Val Arg Glu Glu
65                  70                  75                  80
Val Cys Asp Ile Leu Arg Phe Trp Leu Asp Lys Gly Val Ser Gly Phe
                85                  90                  95
Arg Met Asp Val Ile Asn Leu Ile Ser Lys Asp Gln Arg Phe Pro Asp
            100                 105                 110
Ala Glu Ile Arg His Pro Gly Arg Lys Tyr Gln Pro Gly Glu Cys Tyr
        115                 120                 125
Phe Ala Asn Gly Val Arg Leu Met Asp Tyr Leu Gln Glu Met Lys Thr
    130                 135                 140
Ala Val Phe Ser Lys Tyr Asp Ala Val Thr Val Gly Glu Met Pro Tyr
145                 150                 155                 160
Leu Glu Asp Glu Glu Gln Arg Leu Arg Met Val Ala Ala Glu Glu Gly
                165                 170                 175
Val Leu Asn Met Ile Phe Thr Phe Glu Met Ile Gly Leu Asp Ile Val
            180                 185                 190
Pro Glu Lys Gly Arg Phe Ser Asn Lys Ala Trp Lys Val Cys Glu Leu
        195                 200                 205
Lys Asp Ile Val Ala Lys Ala Cys Lys Leu Val Asn Gln Asp Gly Trp
    210                 215                 220
His Thr Leu Phe Cys Glu Asn His Asp Gln Pro Arg Ser Val Thr Arg
225                 230                 235                 240
Phe Cys Asp Asp Ser Asp Glu His Arg Val Ala Gly Thr Lys Leu Leu
                245                 250                 255
Ser Ile Met Gln Thr Ser Leu Pro Gly Thr Leu Tyr Val Tyr Gln Gly
            260                 265                 270
Glu Glu Leu Gly Met Arg Asn Val Pro Lys Ser Trp Gly Phe Glu Glu
        275                 280                 285
Tyr Ile Asp Ile Glu Ser Val Ser Tyr Ile Lys Asn
    290                 295                 300

<210> 507
<211> 134
<212> PRT
<213> Fungi

<400> 507
Val Cys Ala Gln Phe Pro Glu Gly Ser Pro Glu Arg Lys Glu Ala Lys
 1               5                  10                  15
Asn Leu Leu Arg Leu Lys Ala Arg Asp Asn Ala Arg Thr Pro Met Gln
            20                  25                  30
Trp Asp Ser Thr Ser Asn Gly Gly Phe Cys Pro Glu Gly Val Lys Pro
        35                  40                  45
Trp Met Arg Val Asn Asp Asp Tyr Thr Met Val Asn Ala Ala Leu Gln
    50                  55                  60
Thr Ser Ala Gly Gln Ala Thr Asp Arg Thr Met Leu Val Ser Pro Tyr
65                  70                  75                  80
Arg Phe Trp Gln Arg Ser Ile Gln Ile Arg Lys Lys Phe Lys Asp Leu
                85                  90                  95
Leu Val Tyr Gly Asp Leu Glu Ile Ile Asp Gly Thr His Ala Asn Val
            100                 105                 110
Phe Ala Phe Lys Arg Ile Ser Asn Gly Arg His Thr Ile Thr Ile Leu
        115                 120                 125
Asn Phe Ser Lys Glu Glu
    130

<210> 508
<211> 61
<212> PRT
<213> Fungi
```

<400> 508
Ala Val Asp Gly Gly Asp Thr Ala Phe Thr Leu Gly Gly Leu Ala Trp
 1   5   10   15
Tyr Cys Leu Met Asp Leu Ser Phe Trp Asn Gly Ser Lys Arg Leu Val
            20   25   30
Ser Met Ala Thr Ile Leu Ser Leu Tyr Ser Gln Tyr Arg Met Pro Arg
        35   40   45
Thr Val Ser Asn Met Gln Cys Leu Val Ile Pro Ser Ser
    50   55   60

<210> 509
<211> 1833
<212> DNA
<213> Fungi

<400> 509
atgagtcaag tttctggaaa agtgcccgac aggaagtggt ggaaggaagc tattgtctac 60
cagatctatc ccgcgtcgtt tcttgatacc gattctaatg gcgtcgggaa catcaatggg 120
atcacagcta agctggacta cctcaaggaa ttgggcgtgg acattctctg gatttcaccc 180
gtttacgaaa gcccccagaa ggacatggga tacgacattt cgaattaccg cgatatccat 240
cgcccctacg gcaccatgca agacgtgcac cacttgatca acgagctcaa agcacgcgac 300
atgaaactga tcatggatct ggtggtcaat cacacctcga ccgagcatcc ctggttcatc 360
gactcagcct cgtccgcagt aagcgcgcga cgagactggt acatctggcg caagccaaag 420
tacgacgagc agggcaatcg gcagccaccg aataactggt gcagcctctt tgacgagact 480
gaatcagcct ggacctacga cgctaagacg gatgagtact atctctcgct gtttagtccg 540
ttccaggcag atcttaactg ggagacccca ttcgtcagag aggaagtctg cgacatcctg 600
cgcttctggc tcgacaaggg cgtctctggt ttccggatgg acgtcatcaa cctcatctcc 660
aaggaccaga ggttccctga cgcagagatc cgccacccgg gcaggaagta ccagccgggt 720
gaatgctact tcgccaacgg ggtccgactc atggactatc tccaggagat gaagaccgcc 780
gtattctcaa agtatgacgc cgtcacggtc ggcgagatgc cgtaccttga agatgaagag 840
cagagactca ggatggtcgc ggctgaagaa ggcgtgctga acatgatctt cacttttgag 900
atgattggcc tcgacatcgt ccccgaaaaa ggccgcttca gcaataaagc ttggaaggtc 960
tgtgagctga aggacatagt tgccaaggcc tgcaagctcg ttaaccagga cgggtggcat 1020
acgctcttct gcgagaatca tgaccagccg cgttccgtca cgcgtttctg cgacgactcg 1080
gatgagcatc gtgtcgcagg caccaagctc ctctccatca tgcagacaag ccttcctggt 1140
accctgtatg tctaccaggg cgaagagctg ggtatgcgaa acgtgcccaa gtcctgggga 1200
tttgaagaat acatagacat tgagtctgtc tcctacatca agaatgtctg cgctcaattc 1260
ccagaagggt ctccggaaag aaaagaggct aagaatcttc tgcgcctcaa agctcgtgat 1320
aatgccagga ccccaatgca gtgggactct acatctaacg gcggcttctg tcccgaggga 1380
gtcaagccct ggatgcgcgt caacgacgac tacacgatgg tcaacgcagc tctacagact 1440
tccgcaggcc aggcaaccga ccgcaccatg cttgtgtcac cctataggtt ctggcagcga 1500
agcatccaga tacgcaagaa gttcaaagac ctgctcgtgt acggagacct tgagattatt 1560
gatggcacgc acgccaacgt ctttgcgttc aagaggatca gcaacggcag acataccatc 1620
acgatcctta actttagtaa ggaggaggcc gtggacgggg gagatacagc ttttaccttg 1680
ggagggcttg cttggtattg cttgatggat ctctcatttt ggaatggaag taaaagattg 1740
gtatcaatgg ctacgatatt gtctttatat agccagtaca gaatgccacg gactgttagt 1800
aacatgcaat gtcttgtgat accatccagt taa 1833

<210> 510
<211> 610
<212> PRT
<213> Fungi

<220>
<221> DOMAIN
<222> (21)...(443)
<223> Alpha amylase, catalytic domain

<400> 510
Met Ser Gln Val Ser Gly Lys Val Pro Asp Arg Lys Trp Trp Lys Glu
 1   5   10   15
Ala Ile Val Tyr Gln Ile Tyr Pro Ala Ser Phe Leu Asp Thr Asp Ser
            20   25   30
Asn Gly Val Gly Asn Ile Asn Gly Ile Thr Ala Lys Leu Asp Tyr Leu
        35   40   45
Lys Glu Leu Gly Val Asp Ile Leu Trp Ile Ser Pro Val Tyr Glu Ser
    50   55   60
Pro Gln Lys Asp Met Gly Tyr Asp Ile Ser Asn Tyr Arg Asp Ile His

```
65                  70                  75                  80
Arg Pro Tyr Gly Thr Met Gln Asp Val His His Leu Ile Asn Glu Leu
                85                  90                  95
Lys Ala Arg Asp Met Lys Leu Ile Met Asp Leu Val Val Asn His Thr
            100                 105                 110
Ser Thr Glu His Pro Trp Phe Ile Asp Ser Ala Ser Ser Ala Val Ser
        115                 120                 125
Ala Arg Arg Asp Trp Tyr Ile Trp Arg Lys Pro Lys Tyr Asp Glu Gln
    130                 135                 140
Gly Asn Arg Gln Pro Pro Asn Asn Trp Cys Ser Leu Phe Asp Glu Thr
145                 150                 155                 160
Glu Ser Ala Trp Thr Tyr Asp Ala Lys Thr Asp Glu Tyr Tyr Leu Ser
                165                 170                 175
Leu Phe Ser Pro Phe Gln Ala Asp Leu Asn Trp Glu Thr Pro Phe Val
            180                 185                 190
Arg Glu Glu Val Cys Asp Ile Leu Arg Phe Trp Leu Asp Lys Gly Val
        195                 200                 205
Ser Gly Phe Arg Met Asp Val Ile Asn Leu Ile Ser Lys Asp Gln Arg
    210                 215                 220
Phe Pro Asp Ala Glu Ile Arg His Pro Gly Arg Lys Tyr Gln Pro Gly
225                 230                 235                 240
Glu Cys Tyr Phe Ala Asn Gly Val Arg Leu Met Asp Tyr Leu Gln Glu
                245                 250                 255
Met Lys Thr Ala Val Phe Ser Lys Tyr Asp Ala Val Thr Val Gly Glu
            260                 265                 270
Met Pro Tyr Leu Glu Asp Glu Glu Gln Arg Leu Arg Met Val Ala Ala
        275                 280                 285
Glu Glu Gly Val Leu Asn Met Ile Phe Thr Phe Glu Met Ile Gly Leu
    290                 295                 300
Asp Ile Val Pro Glu Lys Gly Arg Phe Ser Asn Lys Ala Trp Lys Val
305                 310                 315                 320
Cys Glu Leu Lys Asp Ile Val Ala Lys Ala Cys Lys Leu Val Asn Gln
                325                 330                 335
Asp Gly Trp His Thr Leu Phe Cys Glu Asn His Asp Gln Pro Arg Ser
            340                 345                 350
Val Thr Arg Phe Cys Asp Asp Ser Asp Glu His Arg Val Ala Gly Thr
        355                 360                 365
Lys Leu Leu Ser Ile Met Gln Thr Ser Leu Pro Gly Thr Leu Tyr Val
    370                 375                 380
Tyr Gln Gly Glu Glu Leu Gly Met Arg Asn Val Pro Lys Ser Trp Gly
385                 390                 395                 400
Phe Glu Glu Tyr Ile Asp Ile Glu Ser Val Ser Tyr Ile Lys Asn Val
                405                 410                 415
Cys Ala Gln Phe Pro Glu Gly Ser Pro Glu Arg Lys Glu Ala Lys Asn
            420                 425                 430
Leu Leu Arg Leu Lys Ala Arg Asp Asn Ala Arg Thr Pro Met Gln Trp
        435                 440                 445
Asp Ser Thr Ser Asn Gly Gly Phe Cys Pro Glu Gly Val Lys Pro Trp
    450                 455                 460
Met Arg Val Asn Asp Asp Tyr Thr Met Val Asn Ala Ala Leu Gln Thr
465                 470                 475                 480
Ser Ala Gly Gln Ala Thr Asp Arg Thr Met Leu Val Ser Pro Tyr Arg
                485                 490                 495
Phe Trp Gln Arg Ser Ile Gln Ile Arg Lys Lys Phe Lys Asp Leu Leu
            500                 505                 510
Val Tyr Gly Asp Leu Glu Ile Ile Asp Gly Thr His Ala Asn Val Phe
        515                 520                 525
Ala Phe Lys Arg Ile Ser Asn Gly Arg His Thr Ile Thr Ile Leu Asn
    530                 535                 540
Phe Ser Lys Glu Glu Ala Val Asp Gly Gly Asp Thr Ala Phe Thr Leu
545                 550                 555                 560
Gly Gly Leu Ala Trp Tyr Cys Leu Met Asp Leu Ser Phe Trp Asn Gly
                565                 570                 575
Ser Lys Arg Leu Val Ser Met Ala Thr Ile Leu Ser Leu Tyr Ser Gln
            580                 585                 590
Tyr Arg Met Pro Arg Thr Val Ser Asn Met Gln Cys Leu Val Ile Pro
        595                 600                 605
Ser Ser
```

610

<210> 511
<211> 1792
<212> DNA
<213> Fungi

<220>
<221> CDS
<222> (1)...(300)
<223> Exon

<221> CDS
<222> (392)...(1081)
<223> Exon

<221> CDS
<222> (1133)...(1792)
<223> Exon

<400> 511

```
atg aag tac tca att att cca ttc gtt ccc ctt ttc gcg ggt tta tct       48
Met Lys Tyr Ser Ile Ile Pro Phe Val Pro Leu Phe Ala Gly Leu Ser
 1               5                  10                  15

cga gca gca agt agt gat gac tgg gcg ggg agg tct atc tat cag gtt       96
Arg Ala Ala Ser Ser Asp Asp Trp Ala Gly Arg Ser Ile Tyr Gln Val
            20                  25                  30

atc aca gat aga tat cat cga tcc tca gat tcg gat gct cct tgt aat      144
Ile Thr Asp Arg Tyr His Arg Ser Ser Asp Ser Asp Ala Pro Cys Asn
         35                  40                  45

atc acg aat tat tgt gga ggt acc tgg aat ggc atc acg gag aat ctg      192
Ile Thr Asn Tyr Cys Gly Gly Thr Trp Asn Gly Ile Thr Glu Asn Leu
     50                  55                  60

gat tat atc cag aat atg ggg ttc acg gcc att caa att agt ccc gtc      240
Asp Tyr Ile Gln Asn Met Gly Phe Thr Ala Ile Gln Ile Ser Pro Val
 65                  70                  75                  80

aat ttg aac att aac agt acg acg att tac gga caa gct ttc cat gga      288
Asn Leu Asn Ile Asn Ser Thr Thr Ile Tyr Gly Gln Ala Phe His Gly
                85                  90                  95

tat tgg cag act gtaggttttc tcttggcaga gcttccctct atttgcatac          340
Tyr Trp Gln Thr
            100

aattggcttg ttgttggggt agcatgaaga gttctaacat caatactgca g tct ctc     397
                                                      Ser Leu


tac gat ctc aat ccg aac ttc ggt tca gca gat gat ttg ttg aaa cta      445
Tyr Asp Leu Asn Pro Asn Phe Gly Ser Ala Asp Asp Leu Leu Lys Leu
        105                 110                 115

agt gca gaa gtt cac aag aga aaa atg tat ctc cta gtt gat gtc gtt      493
Ser Ala Glu Val His Lys Arg Lys Met Tyr Leu Leu Val Asp Val Val
    120                 125                 130

gct agt gag atg gcc gtt gac att gga gat cac aac atg acc gct gga      541
Ala Ser Glu Met Ala Val Asp Ile Gly Asp His Asn Met Thr Ala Gly
135                 140                 145                 150

acc aag atc gat tat tca gct ttc tcc cct gct cca ttc aac gat gaa      589
Thr Lys Ile Asp Tyr Ser Ala Phe Ser Pro Ala Pro Phe Asn Asp Glu
                155                 160                 165
```

```
tca tca tac aac ccc tac tgc cca att atc gat tgg caa aat gtc aca      637
Ser Ser Tyr Asn Pro Tyr Cys Pro Ile Ile Asp Trp Gln Asn Val Thr
            170                 175                 180

gaa tat caa gat tgt tgg ttg gga tac acc gga gtc gca act cca gat      685
Glu Tyr Gln Asp Cys Trp Leu Gly Tyr Thr Gly Val Ala Thr Pro Asp
        185                 190                 195

atc aaa aca gac aat aaa gaa atc gca gcc gca tta ggc aca tgg att      733
Ile Lys Thr Asp Asn Lys Glu Ile Ala Ala Ala Leu Gly Thr Trp Ile
    200                 205                 210

aag gag ctt gtc tcc aat tac tct att gat gga att cga gtc gat ggt      781
Lys Glu Leu Val Ser Asn Tyr Ser Ile Asp Gly Ile Arg Val Asp Gly
215                 220                 225                 230

gct aaa cag att aca tat gac ttc ttc caa gat ttc gtc gat gcc gcc      829
Ala Lys Gln Ile Thr Tyr Asp Phe Phe Gln Asp Phe Val Asp Ala Ala
                235                 240                 245

ggt gta tat cct ttg gga gaa gta gag gat ggc gat gct gcc ttt act      877
Gly Val Tyr Pro Leu Gly Glu Val Glu Asp Gly Asp Ala Ala Phe Thr
            250                 255                 260

tgc aat tac caa cag tac acc gag gga ttg gag aac tat ccg gtt tat      925
Cys Asn Tyr Gln Gln Tyr Thr Glu Gly Leu Glu Asn Tyr Pro Val Tyr
        265                 270                 275

tac aca acc att caa gcc ttt acc gct gga aaa atg gct ggt ctg gtc      973
Tyr Thr Thr Ile Gln Ala Phe Thr Ala Gly Lys Met Ala Gly Leu Val
    280                 285                 290

gat atg gtc aag agt atg gga gct ttg tgc aaa tct cca caa tat ctc     1021
Asp Met Val Lys Ser Met Gly Ala Leu Cys Lys Ser Pro Gln Tyr Leu
295                 300                 305                 310

gcg aat ttc att gag aat caa gat caa cca cgt ttt gca tct tat aat     1069
Ala Asn Phe Ile Glu Asn Gln Asp Gln Pro Arg Phe Ala Ser Tyr Asn
                315                 320                 325

gat gat gag act gtaagtttct tctacttcat gttctaatac aaaatcacta         1121
Asp Asp Glu Thr
            330

acacaaaaca g ctc gca gcc aac gca atg gct ttc aca atc ctg gca gat    1171
             Leu Ala Ala Asn Ala Met Ala Phe Thr Ile Leu Ala Asp
                         335                 340

ggt atc ccc aag tac tat tac gga caa gaa caa cat ctg cag ggc aac     1219
Gly Ile Pro Lys Tyr Tyr Tyr Gly Gln Glu Gln His Leu Gln Gly Asn
    345                 350                 355

tac tct cca tac aac cgt caa gct ctc tgg gaa agt aag cca agc aca     1267
Tyr Ser Pro Tyr Asn Arg Gln Ala Leu Trp Glu Ser Lys Pro Ser Thr
360                 365                 370                 375

act act cct ctc tac act atg aca gct act ctc aac acc ctc cgt aac     1315
Thr Thr Pro Leu Tyr Thr Met Thr Ala Thr Leu Asn Thr Leu Arg Asn
                380                 385                 390

cac gcc atc tcc ata aac tca aac tac atc gcc aac atg tcc att ctg     1363
His Ala Ile Ser Ile Asn Ser Asn Tyr Ile Ala Asn Met Ser Ile Leu
            395                 400                 405

ctc cac aat gac gat tcc aca tac gca aca cgt aaa gga cca aac ggt     1411
Leu His Asn Asp Asp Ser Thr Tyr Ala Thr Arg Lys Gly Pro Asn Gly
        410                 415                 420
```

```
gtt caa atc gtc tcg gta ctt tct aac caa ggt tcc aaa gga gga agc     1459
Val Gln Ile Val Ser Val Leu Ser Asn Gln Gly Ser Lys Gly Gly Ser
    425                 430                 435

tac act cta ggc atc aac aat gcc gcc gac gct gga aca aac atg acc     1507
Tyr Thr Leu Gly Ile Asn Asn Ala Ala Asp Ala Gly Thr Asn Met Thr
440                 445                 450                 455

gaa gta atc aca tgc aca tcg tac att gca agt gag aac gga acg ctc     1555
Glu Val Ile Thr Cys Thr Ser Tyr Ile Ala Ser Glu Asn Gly Thr Leu
                460                 465                 470

ttc att gag atg gac aaa gga gag cca aga gta ctg ttc cct agc tac     1603
Phe Ile Glu Met Asp Lys Gly Glu Pro Arg Val Leu Phe Pro Ser Tyr
            475                 480                 485

caa atg aat gga acg ggc att tgt gga ttc gat gaa gat gct agt gcg     1651
Gln Met Asn Gly Thr Gly Ile Cys Gly Phe Asp Glu Asp Ala Ser Ala
        490                 495                 500

gga aca gca aca agc act gga acg gcg act gga agt gca gct agt gcg     1699
Gly Thr Ala Thr Ser Thr Gly Thr Ala Thr Gly Ser Ala Ala Ser Ala
    505                 510                 515

aca tca tcg aag aag gga gat gcg agt gga ttg aga gtt gcc ggt tgg     1747
Thr Ser Ser Lys Lys Gly Asp Ala Ser Gly Leu Arg Val Ala Gly Trp
520                 525                 530                 535

gtg ttc gcg ggt gcg atg ggt gcg tct gct ttg gct ttc atg tag         1792
Val Phe Ala Gly Ala Met Gly Ala Ser Ala Leu Ala Phe Met  *
                540                 545
```

<210> 512
<211> 100
<212> PRT
<213> Fungi

<400> 512

```
Met Lys Tyr Ser Ile Ile Pro Phe Val Pro Leu Phe Ala Gly Leu Ser
 1               5                  10                  15
Arg Ala Ala Ser Ser Asp Asp Trp Ala Gly Arg Ser Ile Tyr Gln Val
            20                  25                  30
Ile Thr Asp Arg Tyr His Arg Ser Ser Asp Ser Asp Ala Pro Cys Asn
        35                  40                  45
Ile Thr Asn Tyr Cys Gly Gly Thr Trp Asn Gly Ile Thr Glu Asn Leu
    50                  55                  60
Asp Tyr Ile Gln Asn Met Gly Phe Thr Ala Ile Gln Ile Ser Pro Val
65                  70                  75                  80
Asn Leu Asn Ile Asn Ser Thr Thr Ile Tyr Gly Gln Ala Phe His Gly
                85                  90                  95
Tyr Trp Gln Thr
            100
```

<210> 513
<211> 230
<212> PRT
<213> Fungi

<400> 513

```
Ser Leu Tyr Asp Leu Asn Pro Asn Phe Gly Ser Ala Asp Asp Leu Leu
 1               5                  10                  15
Lys Leu Ser Ala Glu Val His Lys Arg Lys Met Tyr Leu Leu Val Asp
            20                  25                  30
Val Val Ala Ser Glu Met Ala Val Asp Ile Gly Asp His Asn Met Thr
        35                  40                  45
Ala Gly Thr Lys Ile Asp Tyr Ser Ala Phe Ser Pro Ala Pro Phe Asn
```

```
    50                  55                  60
Asp Glu Ser Ser Tyr Asn Pro Tyr Cys Pro Ile Ile Asp Trp Gln Asn
65                  70                  75                  80
Val Thr Glu Tyr Gln Asp Cys Trp Leu Gly Tyr Thr Gly Val Ala Thr
                85                  90                  95
Pro Asp Ile Lys Thr Asp Asn Lys Glu Ile Ala Ala Ala Leu Gly Thr
            100                 105                 110
Trp Ile Lys Glu Leu Val Ser Asn Tyr Ser Ile Asp Gly Ile Arg Val
        115                 120                 125
Asp Gly Ala Lys Gln Ile Thr Tyr Asp Phe Phe Gln Asp Phe Val Asp
    130                 135                 140
Ala Ala Gly Val Tyr Pro Leu Gly Glu Val Glu Asp Gly Asp Ala Ala
145                 150                 155                 160
Phe Thr Cys Asn Tyr Gln Gln Tyr Thr Glu Gly Leu Glu Asn Tyr Pro
                165                 170                 175
Val Tyr Tyr Thr Thr Ile Gln Ala Phe Thr Ala Gly Lys Met Ala Gly
            180                 185                 190
Leu Val Asp Met Val Lys Ser Met Gly Ala Leu Cys Lys Ser Pro Gln
        195                 200                 205
Tyr Leu Ala Asn Phe Ile Glu Asn Gln Asp Gln Pro Arg Phe Ala Ser
    210                 215                 220
Tyr Asn Asp Asp Glu Thr
225                 230

<210> 514
<211> 219
<212> PRT
<213> Fungi

<400> 514
Leu Ala Ala Asn Ala Met Ala Phe Thr Ile Leu Ala Asp Gly Ile Pro
 1                   5                  10                  15
Lys Tyr Tyr Tyr Gly Gln Glu Gln His Leu Gln Gly Asn Tyr Ser Pro
            20                  25                  30
Tyr Asn Arg Gln Ala Leu Trp Glu Ser Lys Pro Ser Thr Thr Thr Pro
        35                  40                  45
Leu Tyr Thr Met Thr Ala Thr Leu Asn Thr Leu Arg Asn His Ala Ile
    50                  55                  60
Ser Ile Asn Ser Asn Tyr Ile Ala Asn Met Ser Ile Leu Leu His Asn
65                  70                  75                  80
Asp Asp Ser Thr Tyr Ala Thr Arg Lys Gly Pro Asn Gly Val Gln Ile
                85                  90                  95
Val Ser Val Leu Ser Asn Gln Gly Ser Lys Gly Gly Ser Tyr Thr Leu
            100                 105                 110
Gly Ile Asn Asn Ala Ala Asp Ala Gly Thr Asn Met Thr Glu Val Ile
        115                 120                 125
Thr Cys Thr Ser Tyr Ile Ala Ser Glu Asn Gly Thr Leu Phe Ile Glu
    130                 135                 140
Met Asp Lys Gly Glu Pro Arg Val Leu Phe Pro Ser Tyr Gln Met Asn
145                 150                 155                 160
Gly Thr Gly Ile Cys Gly Phe Asp Glu Asp Ala Ser Ala Gly Thr Ala
                165                 170                 175
Thr Ser Thr Gly Thr Ala Thr Gly Ser Ala Ala Ser Ala Thr Ser Ser
            180                 185                 190
Lys Lys Gly Asp Ala Ser Gly Leu Arg Val Ala Gly Trp Val Phe Ala
        195                 200                 205
Gly Ala Met Gly Ala Ser Ala Leu Ala Phe Met
    210                 215

<210> 515
<211> 1650
<212> DNA
<213> Fungi

<400> 515
atgaagtact caattattcc attcgttccc cttttcgcgg gtttatctcg agcagcaagt      60
agtgatgact gggcggggag gtctatctat caggttatca cagatagata tcatcgatcc     120
tcagattcgg atgctccttg taatatcacg aattattgtg gaggtacctg gaatggcatc     180
```

```
acggagaatc tggattatat ccagaatatg gggttcacgg ccattcaaat tagtcccgtc     240
aatttgaaca ttaacagtac gacgatttac ggacaagctt tccatggata ttggcagact     300
tctctctacg atctcaatcc gaacttcggt tcagcagatg atttgttgaa actaagtgca     360
gaagttcaca agagaaaaat gtatctccta gttgatgtcg ttgctagtga gatggccgtt     420
gacattggag atcacaacat gaccgctgga accaagatcg attattcagc tttctcccct     480
gctccattca acgatgaatc atcatacaac ccctactgcc caattatcga ttggcaaaat     540
gtcacagaat atcaagattg ttggttggga tacaccggag tcgcaactcc agatatcaaa     600
acagacaata aagaaatcgc agccgcatta ggcacatgga ttaaggagct tgtctccaat     660
tactctattg atggaattcg agtcgatggt gctaaacaga ttacatatga cttcttccaa     720
gatttcgtcg atgccgccgg tgtatatcct ttgggagaag tagaggatgg cgatgctgcc     780
tttacttgca attaccaaca gtacaccgag ggattggaga actatccggt ttattacaca     840
accattcaag cctttaccgc tggaaaaatg gctggtctgg tcgatatggt caagagtatg     900
ggagctttgt gcaaatctcc acaatatctc gcgaatttca ttgagaatca agatcaacca     960
cgttttgcat cttataatga tgatgagact ctcgcagcca acgcaatggc tttcacaatc    1020
ctggcagatg gtatccccaa gtactattac ggacaagaac aacatctgca gggcaactac    1080
tctccataca accgtcaagc tctctgggaa agtaagccaa gcacaactac tcctctctac    1140
actatgacag ctactctcaa caccctccgt aaccacgcca tctccataaa ctcaaactac    1200
atcgccaaca tgtccattct gctccacaat gacgattcca catacgcaac acgtaaagga    1260
ccaaacggtg ttcaaatcgt ctcggtactt tctaaccaag gttccaaagg aggaagctac    1320
actctaggca tcaacaatgc cgccgacgct ggaacaaaca tgaccgaagt aatcacatgc    1380
acatcgtaca ttgcaagtga gaacggaacg ctcttcattg agatggacaa aggagagcca    1440
agagtactgt tccctagcta ccaaatgaat ggaacgggca tttgtggatt cgatgaagat    1500
gctagtgcgg gaacagcaac aagcactgga acggcgactg gaagtgcagc tagtgcgaca    1560
tcatcgaaga agggagatgc gagtggattg agagttgccg gttgggtgtt cgcgggtgcg    1620
atgggtgcgt ctgctttggc tttcatgtag                                     1650

<210> 516
<211> 549
<212> PRT
<213> Fungi

<220>
<221> SIGNAL
<222> (1)...(21)

<221> DOMAIN
<222> (31)...(390)
<223> Alpha amylase, catalytic domain

<400> 516
Met Lys Tyr Ser Ile Ile Pro Phe Val Pro Leu Phe Ala Gly Leu Ser
 1               5                  10                  15
Arg Ala Ala Ser Ser Asp Asp Trp Ala Gly Arg Ser Ile Tyr Gln Val
            20                  25                  30
Ile Thr Asp Arg Tyr His Arg Ser Ser Asp Ser Asp Ala Pro Cys Asn
        35                  40                  45
Ile Thr Asn Tyr Cys Gly Gly Thr Trp Asn Gly Ile Thr Glu Asn Leu
    50                  55                  60
Asp Tyr Ile Gln Asn Met Gly Phe Thr Ala Ile Gln Ile Ser Pro Val
65                  70                  75                  80
Asn Leu Asn Ile Asn Ser Thr Thr Ile Tyr Gly Gln Ala Phe His Gly
                85                  90                  95
Tyr Trp Gln Thr Ser Leu Tyr Asp Leu Asn Pro Asn Phe Gly Ser Ala
            100                 105                 110
Asp Asp Leu Leu Lys Leu Ser Ala Glu Val His Lys Arg Lys Met Tyr
        115                 120                 125
Leu Leu Val Asp Val Val Ala Ser Glu Met Ala Val Asp Ile Gly Asp
    130                 135                 140
His Asn Met Thr Ala Gly Thr Lys Ile Asp Tyr Ser Ala Phe Ser Pro
145                 150                 155                 160
Ala Pro Phe Asn Asp Glu Ser Ser Tyr Asn Pro Tyr Cys Pro Ile Ile
                165                 170                 175
Asp Trp Gln Asn Val Thr Glu Tyr Gln Asp Cys Trp Leu Gly Tyr Thr
            180                 185                 190
Gly Val Ala Thr Pro Asp Ile Lys Thr Asp Asn Lys Glu Ile Ala Ala
        195                 200                 205
Ala Leu Gly Thr Trp Ile Lys Glu Leu Val Ser Asn Tyr Ser Ile Asp
    210                 215                 220
```

```
Gly Ile Arg Val Asp Gly Ala Lys Gln Ile Thr Tyr Asp Phe Phe Gln
225                 230                 235                 240
Asp Phe Val Asp Ala Ala Gly Val Tyr Pro Leu Gly Glu Val Glu Asp
                245                 250                 255
Gly Asp Ala Ala Phe Thr Cys Asn Tyr Gln Gln Tyr Thr Glu Gly Leu
            260                 265                 270
Glu Asn Tyr Pro Val Tyr Tyr Thr Thr Ile Gln Ala Phe Thr Ala Gly
        275                 280                 285
Lys Met Ala Gly Leu Val Asp Met Val Lys Ser Met Gly Ala Leu Cys
    290                 295                 300
Lys Ser Pro Gln Tyr Leu Ala Asn Phe Ile Glu Asn Gln Asp Gln Pro
305                 310                 315                 320
Arg Phe Ala Ser Tyr Asn Asp Asp Glu Thr Leu Ala Ala Asn Ala Met
                325                 330                 335
Ala Phe Thr Ile Leu Ala Asp Gly Ile Pro Lys Tyr Tyr Tyr Gly Gln
            340                 345                 350
Glu Gln His Leu Gln Gly Asn Tyr Ser Pro Tyr Asn Arg Gln Ala Leu
        355                 360                 365
Trp Glu Ser Lys Pro Ser Thr Thr Thr Pro Leu Tyr Thr Met Thr Ala
    370                 375                 380
Thr Leu Asn Thr Leu Arg Asn His Ala Ile Ser Ile Asn Ser Asn Tyr
385                 390                 395                 400
Ile Ala Asn Met Ser Ile Leu Leu His Asn Asp Asp Ser Thr Tyr Ala
                405                 410                 415
Thr Arg Lys Gly Pro Asn Gly Val Gln Ile Val Ser Val Leu Ser Asn
            420                 425                 430
Gln Gly Ser Lys Gly Gly Ser Tyr Thr Leu Gly Ile Asn Asn Ala Ala
        435                 440                 445
Asp Ala Gly Thr Asn Met Thr Glu Val Ile Thr Cys Thr Ser Tyr Ile
    450                 455                 460
Ala Ser Glu Asn Gly Thr Leu Phe Ile Glu Met Asp Lys Gly Glu Pro
465                 470                 475                 480
Arg Val Leu Phe Pro Ser Tyr Gln Met Asn Gly Thr Gly Ile Cys Gly
                485                 490                 495
Phe Asp Glu Asp Ala Ser Ala Gly Thr Ala Thr Ser Thr Gly Thr Ala
            500                 505                 510
Thr Gly Ser Ala Ala Ser Ala Thr Ser Ser Lys Lys Gly Asp Ala Ser
        515                 520                 525
Gly Leu Arg Val Ala Gly Trp Val Phe Ala Gly Ala Met Gly Ala Ser
    530                 535                 540
Ala Leu Ala Phe Met
545

<210> 517
<211> 1758
<212> DNA
<213> Fungi

<220>
<221> CDS
<222> (1)...(1758)
<223> Exon

<400> 517
atg aat atg aac att ttc ctt ttg atc atc tct ctc gct ttt ttc tcc       48
Met Asn Met Asn Ile Phe Leu Leu Ile Ile Ser Leu Ala Phe Phe Ser
 1               5                  10                  15

acc gtc aat tgt tac acc atg tcc aat gcc gag gcc tgg aaa agg cgt       96
Thr Val Asn Cys Tyr Thr Met Ser Asn Ala Glu Ala Trp Lys Arg Arg
            20                  25                  30

tcc ata tac caa gtg ctc acc gac aga ttt gcc ata gaa gat gag aat      144
Ser Ile Tyr Gln Val Leu Thr Asp Arg Phe Ala Ile Glu Asp Glu Asn
        35                  40                  45

aat cct agt aga ggc cca gat cat ggg atg aga gag tat tgt ggt ggg      192
Asn Pro Ser Arg Gly Pro Asp His Gly Met Arg Glu Tyr Cys Gly Gly
```

```
     50                  55                  60

aca tgg agg ggg ata ata tct aaa ctt gat tat att caa gga atg ggt      240
Thr Trp Arg Gly Ile Ile Ser Lys Leu Asp Tyr Ile Gln Gly Met Gly
 65                  70                  75                  80

ttt gat gca gtt tgg att tcg ccc gtt tgt ttc ttc agc ctc acc agc      288
Phe Asp Ala Val Trp Ile Ser Pro Val Cys Phe Phe Ser Leu Thr Ser
                 85                  90                  95

gtt ttg att gca ttt act cac aat tcc ata cag atc acc aag aac atc      336
Val Leu Ile Ala Phe Thr His Asn Ser Ile Gln Ile Thr Lys Asn Ile
            100                 105                 110

gaa gaa ttg act gaa tat gga act gcc tat cac ggt tat tgg cca gag      384
Glu Glu Leu Thr Glu Tyr Gly Thr Ala Tyr His Gly Tyr Trp Pro Glu
        115                 120                 125

gat att tat tcg gta aat cca cat ttt gga acc gct gat gat ttg ctc      432
Asp Ile Tyr Ser Val Asn Pro His Phe Gly Thr Ala Asp Asp Leu Leu
    130                 135                 140

gct cta tcg gat gcc ttg cac gct cgg ggt atg ttt ttg atg gtt gat      480
Ala Leu Ser Asp Ala Leu His Ala Arg Gly Met Phe Leu Met Val Asp
145                 150                 155                 160

ata gtt gtc aat cat atg ggt tcg cct cct aca gtg gac tat tct aga      528
Ile Val Val Asn His Met Gly Ser Pro Pro Thr Val Asp Tyr Ser Arg
                165                 170                 175

tac gta cct ttc aat gat tcg tca tac tat cat cca cag tca ttc gtc      576
Tyr Val Pro Phe Asn Asp Ser Ser Tyr Tyr His Pro Gln Ser Phe Val
            180                 185                 190

acg aac tac gaa aat cag cag gag acc gag gaa gga tgg ttg ggg gat      624
Thr Asn Tyr Glu Asn Gln Gln Glu Thr Glu Glu Gly Trp Leu Gly Asp
        195                 200                 205

gaa cat gtg cca tta ccc gat cta aat acg gag gat cct act gtg gta      672
Glu His Val Pro Leu Pro Asp Leu Asn Thr Glu Asp Pro Thr Val Val
    210                 215                 220

ctc acc ctc caa acc tgg atc aca aaa ctt gta cag agg ttc aaa atc      720
Leu Thr Leu Gln Thr Trp Ile Thr Lys Leu Val Gln Arg Phe Lys Ile
225                 230                 235                 240

gat ggt ttg aga att gac acc ttc aaa cat gtc agg aag aat ttt tgg      768
Asp Gly Leu Arg Ile Asp Thr Phe Lys His Val Arg Lys Asn Phe Trp
                245                 250                 255

ccg gat ttc gtt cgg gct gcg ggg gtc tgg tcg gtg ggt gag gtg ttg      816
Pro Asp Phe Val Arg Ala Ala Gly Val Trp Ser Val Gly Glu Val Leu
            260                 265                 270

tca ggt gac ccc gac tac ttg agc tca tat caa cct tat gtt gga ggc      864
Ser Gly Asp Pro Asp Tyr Leu Ser Ser Tyr Gln Pro Tyr Val Gly Gly
        275                 280                 285

ctt cta gac tat gga aca tat tat cct ttg aag aga gcc ttt cat agg      912
Leu Leu Asp Tyr Gly Thr Tyr Tyr Pro Leu Lys Arg Ala Phe His Arg
    290                 295                 300

gac ggt gga agt atg tat gag ttg acc aac ctc ctc act cct caa tat      960
Asp Gly Gly Ser Met Tyr Glu Leu Thr Asn Leu Leu Thr Pro Gln Tyr
305                 310                 315                 320

cga tca aaa ttc cgc gac atg cag caa atg tcg act ttt atg gag aat     1008
Arg Ser Lys Phe Arg Asp Met Gln Gln Met Ser Thr Phe Met Glu Asn
```

```
                325                 330                 335

cat gat atg ccg cgg ttt acc aac gac acc aac gcg gat ttg gct gtt     1056
His Asp Met Pro Arg Phe Thr Asn Asp Thr Asn Ala Asp Leu Ala Val
            340                 345                 350

gta aag aat gct atg gca tgg aat cta ctc acc gat ggt gtg cca att     1104
Val Lys Asn Ala Met Ala Trp Asn Leu Leu Thr Asp Gly Val Pro Ile
        355                 360                 365

att tac tat gga caa gaa caa tct tat tct gga gat ggt gat cct ggc     1152
Ile Tyr Tyr Gly Gln Glu Gln Ser Tyr Ser Gly Asp Gly Asp Pro Gly
    370                 375                 380

tca cga gag ctg atg tgg aca tcc aat tat aac atc aca cct cta tac     1200
Ser Arg Glu Leu Met Trp Thr Ser Asn Tyr Asn Ile Thr Pro Leu Tyr
385                 390                 395                 400

caa tat atc gca cgt ttg aat cag atc aga aag ctg ggt tgg gat gta     1248
Gln Tyr Ile Ala Arg Leu Asn Gln Ile Arg Lys Leu Gly Trp Asp Val
                405                 410                 415

gga ttt ggt aca cat cta acc aca cgt ctt tac gtt gag gat cat gtc     1296
Gly Phe Gly Thr His Leu Thr Thr Arg Leu Tyr Val Glu Asp His Val
            420                 425                 430

gct gtt acg cag aag ggc ccc gtg ctc atg gtg ctg acc aac cta ggt     1344
Ala Val Thr Gln Lys Gly Pro Val Leu Met Val Leu Thr Asn Leu Gly
        435                 440                 445

agc cag gcc aga tcc aaa act atc tct aca cct acc atg ttc aaa aac     1392
Ser Gln Ala Arg Ser Lys Thr Ile Ser Thr Pro Thr Met Phe Lys Asn
    450                 455                 460

ggt acc ata ata gtc gac att ctc aca gga gaa tcc ttc aga gtc aaa     1440
Gly Thr Ile Ile Val Asp Ile Leu Thr Gly Glu Ser Phe Arg Val Lys
465                 470                 475                 480

cga tcg aca aac atc acc ata gta tct gga gaa cct cga att ttt ctt     1488
Arg Ser Thr Asn Ile Thr Ile Val Ser Gly Glu Pro Arg Ile Phe Leu
                485                 490                 495

cct ctt ggt cta gca gag aat atc tgt gaa cag atc tta cca cca gcg     1536
Pro Leu Gly Leu Ala Glu Asn Ile Cys Glu Gln Ile Leu Pro Pro Ala
            500                 505                 510

cag agt gcc ata tcc aaa cta tta tct aga atg ttt ggt gcc tcc tct     1584
Gln Ser Ala Ile Ser Lys Leu Leu Ser Arg Met Phe Gly Ala Ser Ser
        515                 520                 525

tcc ttg aaa act gcc gcc gaa atc acc gat tgg cca tat ggt gaa ccg     1632
Ser Leu Lys Thr Ala Ala Glu Ile Thr Asp Trp Pro Tyr Gly Glu Pro
    530                 535                 540

gtc aac agc tca gat att ata ctc gaa gaa aaa gat cca gtc gcg gag     1680
Val Asn Ser Ser Asp Ile Ile Leu Glu Glu Lys Asp Pro Val Ala Glu
545                 550                 555                 560

cag gca gga gcc aag aca ttg atg ccg agc gtt caa gag agt tca tcc     1728
Gln Ala Gly Ala Lys Thr Leu Met Pro Ser Val Gln Glu Ser Ser Ser
                565                 570                 575

ccg tgg tct ata ttc tat acg gcc aga taa                             1758
Pro Trp Ser Ile Phe Tyr Thr Ala Arg  *
            580                 585
```

<210> 518

<211> 585
<212> PRT
<213> Fungi

<220>
<221> DOMAIN
<222> (36)...(410)
<223> Alpha amylase, catalytic domain

<221> SIGNAL
<222> (1)...(26)

<400> 518

```
Met Asn Met Asn Ile Phe Leu Leu Ile Ile Ser Leu Ala Phe Phe Ser
 1               5                  10                  15
Thr Val Asn Cys Tyr Thr Met Ser Asn Ala Glu Ala Trp Lys Arg Arg
            20                  25                  30
Ser Ile Tyr Gln Val Leu Thr Asp Arg Phe Ala Ile Glu Asp Glu Asn
        35                  40                  45
Asn Pro Ser Arg Gly Pro Asp His Gly Met Arg Glu Tyr Cys Gly Gly
    50                  55                  60
Thr Trp Arg Gly Ile Ile Ser Lys Leu Asp Tyr Ile Gln Gly Met Gly
65                  70                  75                  80
Phe Asp Ala Val Trp Ile Ser Pro Val Cys Phe Phe Ser Leu Thr Ser
                85                  90                  95
Val Leu Ile Ala Phe Thr His Asn Ser Ile Gln Ile Thr Lys Asn Ile
            100                 105                 110
Glu Glu Leu Thr Glu Tyr Gly Thr Ala Tyr His Gly Tyr Trp Pro Glu
        115                 120                 125
Asp Ile Tyr Ser Val Asn Pro His Phe Gly Thr Ala Asp Asp Leu Leu
    130                 135                 140
Ala Leu Ser Asp Ala Leu His Ala Arg Gly Met Phe Leu Met Val Asp
145                 150                 155                 160
Ile Val Val Asn His Met Gly Ser Pro Pro Thr Val Asp Tyr Ser Arg
                165                 170                 175
Tyr Val Pro Phe Asn Asp Ser Ser Tyr Tyr His Pro Gln Ser Phe Val
            180                 185                 190
Thr Asn Tyr Glu Asn Gln Gln Glu Thr Glu Glu Gly Trp Leu Gly Asp
        195                 200                 205
Glu His Val Pro Leu Pro Asp Leu Asn Thr Glu Asp Pro Thr Val Val
    210                 215                 220
Leu Thr Leu Gln Thr Trp Ile Thr Lys Leu Val Gln Arg Phe Lys Ile
225                 230                 235                 240
Asp Gly Leu Arg Ile Asp Thr Phe Lys His Val Arg Lys Asn Phe Trp
                245                 250                 255
Pro Asp Phe Val Arg Ala Ala Gly Val Trp Ser Val Gly Glu Val Leu
            260                 265                 270
Ser Gly Asp Pro Asp Tyr Leu Ser Ser Tyr Gln Pro Tyr Val Gly Gly
        275                 280                 285
Leu Leu Asp Tyr Gly Thr Tyr Tyr Pro Leu Lys Arg Ala Phe His Arg
    290                 295                 300
Asp Gly Gly Ser Met Tyr Glu Leu Thr Asn Leu Leu Thr Pro Gln Tyr
305                 310                 315                 320
Arg Ser Lys Phe Arg Asp Met Gln Gln Met Ser Thr Phe Met Glu Asn
                325                 330                 335
His Asp Met Pro Arg Phe Thr Asn Asp Thr Asn Ala Asp Leu Ala Val
            340                 345                 350
Val Lys Asn Ala Met Ala Trp Asn Leu Leu Thr Asp Gly Val Pro Ile
        355                 360                 365
Ile Tyr Tyr Gly Gln Glu Gln Ser Tyr Ser Gly Asp Gly Asp Pro Gly
    370                 375                 380
Ser Arg Glu Leu Met Trp Thr Ser Asn Tyr Asn Ile Thr Pro Leu Tyr
385                 390                 395                 400
Gln Tyr Ile Ala Arg Leu Asn Gln Ile Arg Lys Leu Gly Trp Asp Val
                405                 410                 415
Gly Phe Gly Thr His Leu Thr Thr Arg Leu Tyr Val Glu Asp His Val
            420                 425                 430
Ala Val Thr Gln Lys Gly Pro Val Leu Met Val Leu Thr Asn Leu Gly
```

```
        435                 440                 445
Ser Gln Ala Arg Ser Lys Thr Ile Ser Thr Pro Thr Met Phe Lys Asn
    450                 455                 460
Gly Thr Ile Ile Val Asp Ile Leu Thr Gly Glu Ser Phe Arg Val Lys
465                 470                 475                 480
Arg Ser Thr Asn Ile Thr Ile Val Ser Gly Glu Pro Arg Ile Phe Leu
                485                 490                 495
Pro Leu Gly Leu Ala Glu Asn Ile Cys Glu Gln Ile Leu Pro Pro Ala
            500                 505                 510
Gln Ser Ala Ile Ser Lys Leu Leu Ser Arg Met Phe Gly Ala Ser Ser
        515                 520                 525
Ser Leu Lys Thr Ala Ala Glu Ile Thr Asp Trp Pro Tyr Gly Glu Pro
    530                 535                 540
Val Asn Ser Ser Asp Ile Ile Leu Glu Glu Lys Asp Pro Val Ala Glu
545                 550                 555                 560
Gln Ala Gly Ala Lys Thr Leu Met Pro Ser Val Gln Glu Ser Ser Ser
                565                 570                 575
Pro Trp Ser Ile Phe Tyr Thr Ala Arg
            580                 585

<210> 519
<211> 1843
<212> DNA
<213> Fungi

<220>
<221> CDS
<222> (1)...(887)
<223> Exon

<221> CDS
<222> (943)...(1843)
<223> Exon

<400> 519
atg cag cct gtc atc cca aac gtc gag cga gcc tcg tgg cgg gaa gcc       48
Met Gln Pro Val Ile Pro Asn Val Glu Arg Ala Ser Trp Arg Glu Ala
 1               5                  10                  15

tct gta tat cag atc tac cca cct tct ttt aaa gat acc aat ggg gat       96
Ser Val Tyr Gln Ile Tyr Pro Pro Ser Phe Lys Asp Thr Asn Gly Asp
             20                  25                  30

ggt ata gga gat att caa gga att atc tca gag ctt gac tat atc aag      144
Gly Ile Gly Asp Ile Gln Gly Ile Ile Ser Glu Leu Asp Tyr Ile Lys
         35                  40                  45

tct ctt ggc gtt gat atg gta tgg ctg tct ccc att ttg gca tct ccg      192
Ser Leu Gly Val Asp Met Val Trp Leu Ser Pro Ile Leu Ala Ser Pro
     50                  55                  60

cag gtt gat atg ggt tac gat atc tct gac tat aag aat atc tat ccg      240
Gln Val Asp Met Gly Tyr Asp Ile Ser Asp Tyr Lys Asn Ile Tyr Pro
 65                  70                  75                  80

cct tat ggt acc atg gcc gat cat gat gct tta atc aaa ggc ctc cat      288
Pro Tyr Gly Thr Met Ala Asp His Asp Ala Leu Ile Lys Gly Leu His
                 85                  90                  95

gac cgc ggt ttg aag tat atc ctt gac ctt gtt gtg aac cat aca tcg      336
Asp Arg Gly Leu Lys Tyr Ile Leu Asp Leu Val Val Asn His Thr Ser
            100                 105                 110

gat caa cat cct tgg ttc aaa gag tcc aag tct tca aag aca agc aga      384
Asp Gln His Pro Trp Phe Lys Glu Ser Lys Ser Ser Lys Thr Ser Arg
        115                 120                 125

tat cgc gac tgg tat ttc tgg cgg ccg gct cgc tgg gat ccc gaa acc      432
```

```
Tyr Arg Asp Trp Tyr Phe Trp Arg Pro Ala Arg Trp Asp Pro Glu Thr
    130                 135                 140

ggt aaa cgg atg cca cct aac aat tgg gaa tcc ttt ttt agt aca agc      480
Gly Lys Arg Met Pro Pro Asn Asn Trp Glu Ser Phe Phe Ser Thr Ser
145                 150                 155                 160

gct tgg act tgg gac gct acc aca caa gaa tat tat ctt cac ctt tgg      528
Ala Trp Thr Trp Asp Ala Thr Thr Gln Glu Tyr Tyr Leu His Leu Trp
                165                 170                 175

tca tct gga cag cct gac ttg aac tgg gaa aat ccc gat gtc gtc aat      576
Ser Ser Gly Gln Pro Asp Leu Asn Trp Glu Asn Pro Asp Val Val Asn
            180                 185                 190

gct gtg cac gat atc gta aga ttc tgg ctt gat cgt ggc gtt gat ggg      624
Ala Val His Asp Ile Val Arg Phe Trp Leu Asp Arg Gly Val Asp Gly
        195                 200                 205

ttt cga atg gat gtc att aac tat att tcg aaa aca cct ggc ctg ccc      672
Phe Arg Met Asp Val Ile Asn Tyr Ile Ser Lys Thr Pro Gly Leu Pro
    210                 215                 220

gac gca aag atc aag aaa ccg gga ttt ttt cag aag ccc acc gag cat      720
Asp Ala Lys Ile Lys Lys Pro Gly Phe Phe Gln Lys Pro Thr Glu His
225                 230                 235                 240

tgt gct tgt gga cca cgg ctg cac gaa tac cta cgt ggg att gga tcg      768
Cys Ala Cys Gly Pro Arg Leu His Glu Tyr Leu Arg Gly Ile Gly Ser
                245                 250                 255

ata ctc cag gaa tac gat gca ttt agt gtc ggc gag atg cca gat gca      816
Ile Leu Gln Glu Tyr Asp Ala Phe Ser Val Gly Glu Met Pro Asp Ala
            260                 265                 270

gat ccc gaa gaa gtt ctc aaa gct gtg ggc cag gat cgt gga gaa tta      864
Asp Pro Glu Glu Val Leu Lys Ala Val Gly Gln Asp Arg Gly Glu Leu
        275                 280                 285

gct atg gcg ttt cac ttt gac at gtaagtcata acagaatata gtcttttaca      917
Ala Met Ala Phe His Phe Asp Ile
    290                 295

ggatctcggg ctaacaatta cacag a gat ggg ctt gac cta ggt gca aat cat    970
                             Asp Gly Leu Asp Leu Gly Ala Asn His
                                         300                 305

aga ttt gac ccc gcc gtg ttc caa cca aca gca tta aag aac gta gtc     1018
Arg Phe Asp Pro Ala Val Phe Gln Pro Thr Ala Leu Lys Asn Val Val
                310                 315                 320

aac aaa tgg caa aca ttc atg gca tcc aac gca ggt tgg aac gcc ctg     1066
Asn Lys Trp Gln Thr Phe Met Ala Ser Asn Ala Gly Trp Asn Ala Leu
            325                 330                 335

cat atg gag aac cac gac gag agc cga aca gtt agt cgt tac gct tgc     1114
His Met Glu Asn His Asp Glu Ser Arg Thr Val Ser Arg Tyr Ala Cys
        340                 345                 350

gac tct ccc gcc atg cga aca ata tca gcc aag atg cta gca aac cac     1162
Asp Ser Pro Ala Met Arg Thr Ile Ser Ala Lys Met Leu Ala Asn His
    355                 360                 365

ctt gcc ttc cag tct gga aca ctt ttc atc tat cag gga cag gaa ctt     1210
Leu Ala Phe Gln Ser Gly Thr Leu Phe Ile Tyr Gln Gly Gln Glu Leu
370                 375                 380                 385

gca atg gtt aat ctt ccc cga gaa tgg gga atg gag aaa tac aaa gac     1258
```

```
Ala Met Val Asn Leu Pro Arg Glu Trp Gly Met Glu Lys Tyr Lys Asp
                390                 395                 400

atc gaa tgt ctg aat cac tgg gag ttg gta cag cgt tgt tac cac gac     1306
Ile Glu Cys Leu Asn His Trp Glu Leu Val Gln Arg Cys Tyr His Asp
            405                 410                 415

gat gta aaa aag cag aag atg tat agg gac aga tac cga cag gtc gga     1354
Asp Val Lys Lys Gln Lys Met Tyr Arg Asp Arg Tyr Arg Gln Val Gly
        420                 425                 430

cgt gac aac gct cgc acg ccc atg cag tgg aac tcg gag agt cca tat     1402
Arg Asp Asn Ala Arg Thr Pro Met Gln Trp Asn Ser Glu Ser Pro Tyr
    435                 440                 445

gct gga ttc gtg ccg att ggt tct aag caa gtt gag ccc tgg atg tct     1450
Ala Gly Phe Val Pro Ile Gly Ser Lys Gln Val Glu Pro Trp Met Ser
450                 455                 460                 465

atc cac cct gac ttt ggg act tgg aac gtt tcc acg atg tta gcc gat     1498
Ile His Pro Asp Phe Gly Thr Trp Asn Val Ser Thr Met Leu Ala Asp
                470                 475                 480

tcg cag tcc agc ctt cac cat tgg cgc cgg gtg ctc aag ttt cga aag     1546
Ser Gln Ser Ser Leu His His Trp Arg Arg Val Leu Lys Phe Arg Lys
            485                 490                 495

cag cat tcc aac ata ttt gtt tac gga gga ttc gaa atg cta aat ata     1594
Gln His Ser Asn Ile Phe Val Tyr Gly Gly Phe Glu Met Leu Asn Ile
        500                 505                 510

gaa atc gaa gaa gac gtc att gcg tat ata cgc act gac aac act act     1642
Glu Ile Glu Glu Asp Val Ile Ala Tyr Ile Arg Thr Asp Asn Thr Thr
    515                 520                 525

ggt ggg agc cta act gga ccg acg gag gcg ctg gtt gtg acc agt ttt     1690
Gly Gly Ser Leu Thr Gly Pro Thr Glu Ala Leu Val Val Thr Ser Phe
530                 535                 540                 545

agt gcc act gat atc tgg tgg acc att cct ccc aag gca atg act att     1738
Ser Ala Thr Asp Ile Trp Trp Thr Ile Pro Pro Lys Ala Met Thr Ile
                550                 555                 560

cta ctc ggt gcg tcc aca gat tca aga aag ccc aat atc aat gtg acg     1786
Leu Leu Gly Ala Ser Thr Asp Ser Arg Lys Pro Asn Ile Asn Val Thr
            565                 570                 575

ggg atg gta acc gct ttg ggc aat tat gag ggt gtg aac gaa ttg atg     1834
Gly Met Val Thr Ala Leu Gly Asn Tyr Glu Gly Val Asn Glu Leu Met
        580                 585                 590

gtc aat tag                                                        1843
Val Asn  *
    595


<210> 520
<211> 296
<212> PRT
<213> Fungi

<400> 520
Met Gln Pro Val Ile Pro Asn Val Glu Arg Ala Ser Trp Arg Glu Ala
 1               5                  10                  15
Ser Val Tyr Gln Ile Tyr Pro Pro Ser Phe Lys Asp Thr Asn Gly Asp
            20                  25                  30
Gly Ile Gly Asp Ile Gln Gly Ile Ile Ser Glu Leu Asp Tyr Ile Lys
        35                  40                  45
```

```
Ser Leu Gly Val Asp Met Val Trp Leu Ser Pro Ile Leu Ala Ser Pro
    50                  55                  60
Gln Val Asp Met Gly Tyr Asp Ile Ser Asp Tyr Lys Asn Ile Tyr Pro
65                  70                  75                  80
Pro Tyr Gly Thr Met Ala Asp His Asp Ala Leu Ile Lys Gly Leu His
                85                  90                  95
Asp Arg Gly Leu Lys Tyr Ile Leu Asp Leu Val Val Asn His Thr Ser
            100                 105                 110
Asp Gln His Pro Trp Phe Lys Glu Ser Lys Ser Ser Lys Thr Ser Arg
        115                 120                 125
Tyr Arg Asp Trp Tyr Phe Trp Arg Pro Ala Arg Trp Asp Pro Glu Thr
    130                 135                 140
Gly Lys Arg Met Pro Pro Asn Asn Trp Glu Ser Phe Phe Ser Thr Ser
145                 150                 155                 160
Ala Trp Thr Trp Asp Ala Thr Thr Gln Glu Tyr Tyr Leu His Leu Trp
                165                 170                 175
Ser Ser Gly Gln Pro Asp Leu Asn Trp Glu Asn Pro Asp Val Val Asn
            180                 185                 190
Ala Val His Asp Ile Val Arg Phe Trp Leu Asp Arg Gly Val Asp Gly
        195                 200                 205
Phe Arg Met Asp Val Ile Asn Tyr Ile Ser Lys Thr Pro Gly Leu Pro
    210                 215                 220
Asp Ala Lys Ile Lys Lys Pro Gly Phe Phe Gln Lys Pro Thr Glu His
225                 230                 235                 240
Cys Ala Cys Gly Pro Arg Leu His Glu Tyr Leu Arg Gly Ile Gly Ser
                245                 250                 255
Ile Leu Gln Glu Tyr Asp Ala Phe Ser Val Gly Glu Met Pro Asp Ala
            260                 265                 270
Asp Pro Glu Glu Val Leu Lys Ala Val Gly Gln Asp Arg Gly Glu Leu
        275                 280                 285
Ala Met Ala Phe His Phe Asp Ile
    290                 295

<210> 521
<211> 299
<212> PRT
<213> Fungi

<400> 521
Asp Gly Leu Asp Leu Gly Ala Asn His Arg Phe Asp Pro Ala Val Phe
 1               5                  10                  15
Gln Pro Thr Ala Leu Lys Asn Val Val Asn Lys Trp Gln Thr Phe Met
            20                  25                  30
Ala Ser Asn Ala Gly Trp Asn Ala Leu His Met Glu Asn His Asp Glu
        35                  40                  45
Ser Arg Thr Val Ser Arg Tyr Ala Cys Asp Ser Pro Ala Met Arg Thr
    50                  55                  60
Ile Ser Ala Lys Met Leu Ala Asn His Leu Ala Phe Gln Ser Gly Thr
65                  70                  75                  80
Leu Phe Ile Tyr Gln Gly Gln Glu Leu Ala Met Val Asn Leu Pro Arg
                85                  90                  95
Glu Trp Gly Met Glu Lys Tyr Lys Asp Ile Glu Cys Leu Asn His Trp
            100                 105                 110
Glu Leu Val Gln Arg Cys Tyr His Asp Asp Val Lys Lys Gln Lys Met
        115                 120                 125
Tyr Arg Asp Arg Tyr Arg Gln Val Gly Arg Asp Asn Ala Arg Thr Pro
    130                 135                 140
Met Gln Trp Asn Ser Glu Ser Pro Tyr Ala Gly Phe Val Pro Ile Gly
145                 150                 155                 160
Ser Lys Gln Val Glu Pro Trp Met Ser Ile His Pro Asp Phe Gly Thr
                165                 170                 175
Trp Asn Val Ser Thr Met Leu Ala Asp Ser Gln Ser Ser Leu His His
            180                 185                 190
Trp Arg Arg Val Leu Lys Phe Arg Lys Gln His Ser Asn Ile Phe Val
        195                 200                 205
Tyr Gly Gly Phe Glu Met Leu Asn Ile Glu Ile Glu Glu Asp Val Ile
    210                 215                 220
Ala Tyr Ile Arg Thr Asp Asn Thr Thr Gly Gly Ser Leu Thr Gly Pro
```

225 230 235 240
Thr Glu Ala Leu Val Val Thr Ser Phe Ser Ala Thr Asp Ile Trp Trp
245 250 255
Thr Ile Pro Pro Lys Ala Met Thr Ile Leu Leu Gly Ala Ser Thr Asp
260 265 270
Ser Arg Lys Pro Asn Ile Asn Val Thr Gly Met Val Thr Ala Leu Gly
275 280 285
Asn Tyr Glu Gly Val Asn Glu Leu Met Val Asn
290 295

<210> 522
<211> 1788
<212> DNA
<213> Fungi

<400> 522
atgcagcctg tcatcccaaa cgtcgagcga gcctcgtggc gggaagcctc tgtatatcag 60
atctacccac cttctttaa agataccaat ggggatggta taggagatat tcaaggaatt 120
atctcagagc ttgactatat caagtctctt ggcgttgata tggtatggct gtctcccatt 180
ttggcatctc cgcaggttga tatgggttac gatatctctg actataagaa tatctatccg 240
ccttatggta ccatggccga tcatgatgct ttaatcaaag gcctccatga ccgcggtttg 300
aagtatatcc ttgaccttgt tgtgaaccat acatcggatc aacatccttg gttcaaagag 360
tccaagtctt caaagacaag cagatatcgc gactggtatt tctggcggcc ggctcgctgg 420
gatcccgaaa ccggtaaacg gatgccacct aacaattggg aatccttttt tagtacaagc 480
gcttggactt gggacgctac cacacaagaa tattatcttc acctttggtc atctggacag 540
cctgacttga actgggaaaa tcccgatgtc gtcaatgctg tgcacgatat cgtaagattc 600
tggcttgatc gtggcgttga tgggtttcga atggatgtca ttaactatat ttcgaaaaca 660
cctggcctgc ccgacgcaaa gatcaagaaa ccgggatttt ttcagaagcc caccgagcat 720
tgtgcttgtg gaccacggct gcacgaatac ctacgtggga ttggatcgat actccaggaa 780
tacgatgcat ttagtgtcgg cgagatgcca gatgcagatc ccgaagaagt tctcaaagct 840
gtgggccagg atcgtggaga attagctatg gcgtttcact ttgacataga tgggcttgac 900
ctaggtgcaa atcatagatt tgaccccgcc gtgttccaac caacagcatt aaagaacgta 960
gtcaacaaat ggcaaacatt catggcatcc aacgcaggtt ggaacgccct gcatatggag 1020
aaccacgacg agagccgaac agttagtcgt tacgcttgcg actctcccgc catgcgaaca 1080
atatcagcca agatgctagc aaaccacctt gccttccagt ctggaacact tttcatctat 1140
cagggacagg aacttgcaat ggttaatctt ccccgagaat ggggaatgga gaaatacaaa 1200
gacatcgaat gtctgaatca ctgggagttg gtacagcgtt gttaccacga cgatgtaaaa 1260
aagcagaaga tgtataggga cagataccga caggtcggac gtgacaacgc tcgcacgccc 1320
atgcagtgga actcggagag tccatatgct ggattcgtgc cgattggttc taagcaagtt 1380
gagccctgga tgtctatcca ccctgacttt gggacttgga acgtttccac gatgttagcc 1440
gattcgcagt ccagccttca ccattggcgc cgggtgctca agtttcgaaa gcagcattcc 1500
aacatatttg tttacggagg attcgaaatg ctaaatatag aaatcgaaga agacgtcatt 1560
gcgtatatac gcactgacaa cactactggt gggagcctaa ctggaccgac ggaggcgctg 1620
gttgtgacca gttttagtgc cactgatatc tggtggacca ttcctcccaa ggcaatgact 1680
attctactcg gtgcgtccac agattcaaga aagcccaata tcaatgtgac ggggatggta 1740
accgctttgg gcaattatga gggtgtgaac gaattgatgg tcaattag 1788

<210> 523
<211> 595
<212> PRT
<213> Fungi

<220>
<221> DOMAIN
<222> (20)...(430)
<223> Alpha amylase, catalytic domain

<400> 523
Met Gln Pro Val Ile Pro Asn Val Glu Arg Ala Ser Trp Arg Glu Ala
1 5 10 15
Ser Val Tyr Gln Ile Tyr Pro Pro Ser Phe Lys Asp Thr Asn Gly Asp
20 25 30
Gly Ile Gly Asp Ile Gln Gly Ile Ile Ser Glu Leu Asp Tyr Ile Lys
35 40 45
Ser Leu Gly Val Asp Met Val Trp Leu Ser Pro Ile Leu Ala Ser Pro
50 55 60
Gln Val Asp Met Gly Tyr Asp Ile Ser Asp Tyr Lys Asn Ile Tyr Pro
65 70 75 80

```
Pro Tyr Gly Thr Met Ala Asp His Asp Ala Leu Ile Lys Gly Leu His
                85                  90                  95
Asp Arg Gly Leu Lys Tyr Ile Leu Asp Leu Val Val Asn His Thr Ser
            100                 105                 110
Asp Gln His Pro Trp Phe Lys Glu Ser Lys Ser Ser Lys Thr Ser Arg
        115                 120                 125
Tyr Arg Asp Trp Tyr Phe Trp Arg Pro Ala Arg Trp Asp Pro Glu Thr
    130                 135                 140
Gly Lys Arg Met Pro Pro Asn Asn Trp Glu Ser Phe Phe Ser Thr Ser
145                 150                 155                 160
Ala Trp Thr Trp Asp Ala Thr Thr Gln Glu Tyr Tyr Leu His Leu Trp
                165                 170                 175
Ser Ser Gly Gln Pro Asp Leu Asn Trp Glu Asn Pro Asp Val Val Asn
            180                 185                 190
Ala Val His Asp Ile Val Arg Phe Trp Leu Asp Arg Gly Val Asp Gly
        195                 200                 205
Phe Arg Met Asp Val Ile Asn Tyr Ile Ser Lys Thr Pro Gly Leu Pro
    210                 215                 220
Asp Ala Lys Ile Lys Lys Pro Gly Phe Phe Gln Lys Pro Thr Glu His
225                 230                 235                 240
Cys Ala Cys Gly Pro Arg Leu His Glu Tyr Leu Arg Gly Ile Gly Ser
                245                 250                 255
Ile Leu Gln Glu Tyr Asp Ala Phe Ser Val Gly Glu Met Pro Asp Ala
            260                 265                 270
Asp Pro Glu Glu Val Leu Lys Ala Val Gly Gln Asp Arg Gly Glu Leu
        275                 280                 285
Ala Met Ala Phe His Phe Asp Ile Asp Gly Leu Asp Leu Gly Ala Asn
    290                 295                 300
His Arg Phe Asp Pro Ala Val Phe Gln Pro Thr Ala Leu Lys Asn Val
305                 310                 315                 320
Val Asn Lys Trp Gln Thr Phe Met Ala Ser Asn Ala Gly Trp Asn Ala
                325                 330                 335
Leu His Met Glu Asn His Asp Glu Ser Arg Thr Val Ser Arg Tyr Ala
            340                 345                 350
Cys Asp Ser Pro Ala Met Arg Thr Ile Ser Ala Lys Met Leu Ala Asn
        355                 360                 365
His Leu Ala Phe Gln Ser Gly Thr Leu Phe Ile Tyr Gln Gly Gln Glu
    370                 375                 380
Leu Ala Met Val Asn Leu Pro Arg Glu Trp Gly Met Glu Lys Tyr Lys
385                 390                 395                 400
Asp Ile Glu Cys Leu Asn His Trp Glu Leu Val Gln Arg Cys Tyr His
                405                 410                 415
Asp Asp Val Lys Lys Gln Lys Met Tyr Arg Asp Arg Tyr Arg Gln Val
            420                 425                 430
Gly Arg Asp Asn Ala Arg Thr Pro Met Gln Trp Asn Ser Glu Ser Pro
        435                 440                 445
Tyr Ala Gly Phe Val Pro Ile Gly Ser Lys Gln Val Glu Pro Trp Met
    450                 455                 460
Ser Ile His Pro Asp Phe Gly Thr Trp Asn Val Ser Thr Met Leu Ala
465                 470                 475                 480
Asp Ser Gln Ser Ser Leu His His Trp Arg Arg Val Leu Lys Phe Arg
                485                 490                 495
Lys Gln His Ser Asn Ile Phe Val Tyr Gly Gly Phe Glu Met Leu Asn
            500                 505                 510
Ile Glu Ile Glu Glu Asp Val Ile Ala Tyr Ile Arg Thr Asp Asn Thr
        515                 520                 525
Thr Gly Gly Ser Leu Thr Gly Pro Thr Glu Ala Leu Val Val Thr Ser
    530                 535                 540
Phe Ser Ala Thr Asp Ile Trp Trp Thr Ile Pro Pro Lys Ala Met Thr
545                 550                 555                 560
Ile Leu Leu Gly Ala Ser Thr Asp Ser Arg Lys Pro Asn Ile Asn Val
                565                 570                 575
Thr Gly Met Val Thr Ala Leu Gly Asn Tyr Glu Gly Val Asn Glu Leu
            580                 585                 590
Met Val Asn
        595
```

<210> 524

<211> 1820
<212> DNA
<213> Fungi

<220>
<221> CDS
<222> (1)...(277)
<223> Exon

<221> CDS
<222> (343)...(1820)
<223> Exon

<221> misc_feature
<222> (791)...(800)
<223> n = a, g, c or t

<400> 524

```
atg gga tca ata gcc aga aga acc att ggt tca act cct caa tgg tgg       48
Met Gly Ser Ile Ala Arg Arg Thr Ile Gly Ser Thr Pro Gln Trp Trp
 1               5                  10                  15

aaa gaa gca gtc gta tat caa atc tat ccc gct tcg tat ctc gac acc       96
Lys Glu Ala Val Val Tyr Gln Ile Tyr Pro Ala Ser Tyr Leu Asp Thr
            20                  25                  30

acc gga tct gga gat ggt gat ctc aat ggt atc acc tcg aaa ttg cca      144
Thr Gly Ser Gly Asp Gly Asp Leu Asn Gly Ile Thr Ser Lys Leu Pro
        35                  40                  45

tat atc aga tct ctc ggt gta gat gta gta tgg att tca cct atc tat      192
Tyr Ile Arg Ser Leu Gly Val Asp Val Val Trp Ile Ser Pro Ile Tyr
    50                  55                  60

gca tcc cct atg aat gat atg ggc tat gat att tcg gat tat cga gct      240
Ala Ser Pro Met Asn Asp Met Gly Tyr Asp Ile Ser Asp Tyr Arg Ala
65                  70                  75                  80

atc aac cct atg ttt ggt act atg gaa gat tgg gaa c gtgagtcacc          287
Ile Asn Pro Met Phe Gly Thr Met Glu Asp Trp Glu
                85                  90

tcttcctcat tcttttctca aagctctcaa aaaaaaaaaa attaacagcg cctag gt       344
                                                              Arg


ctt tgc gcc aga gca cac gaa cta ggt ttg aaa ctc gtg atg gat ctc       392
Leu Cys Ala Arg Ala His Glu Leu Gly Leu Lys Leu Val Met Asp Leu
    95                  100                 105

gtg gta aat cac act agt tct gaa cac ccc tgg ttc aaa gaa tcc gtc       440
Val Val Asn His Thr Ser Ser Glu His Pro Trp Phe Lys Glu Ser Val
110                 115                 120                 125

tcc ggc ggt ccc aac ggt cct aaa aga gac ttc tac tac tgg caa ccg       488
Ser Gly Gly Pro Asn Gly Pro Lys Arg Asp Phe Tyr Tyr Trp Gln Pro
                130                 135                 140

ccc aaa aat ggt aaa gag ccc aat aac tgg ggt gcc atg ttt gga ggt       536
Pro Lys Asn Gly Lys Glu Pro Asn Asn Trp Gly Ala Met Phe Gly Gly
            145                 150                 155

tca tcc tgg gaa aaa gat cct tcg cat caa aca gac gaa tac tat ctc       584
Ser Ser Trp Glu Lys Asp Pro Ser His Gln Thr Asp Glu Tyr Tyr Leu
        160                 165                 170

cac gtc tac gac gtc agc cag ccc gat ctt aat tgg aca aat cca gcc       632
His Val Tyr Asp Val Ser Gln Pro Asp Leu Asn Trp Thr Asn Pro Ala
```

```
    175                 180                 185

gtt cgc aac gag gtc tgg gat atc atg cgc ttt tgg ctt gat aag gga      680
Val Arg Asn Glu Val Trp Asp Ile Met Arg Phe Trp Leu Asp Lys Gly
190                 195                 200                 205

tgt gat ggg ttt cgt atg gat gtc atc aat tgc att tca aaa gaa cct      728
Cys Asp Gly Phe Arg Met Asp Val Ile Asn Cys Ile Ser Lys Glu Pro
                210                 215                 220

gga ttt ccc gat ttt gaa gtc acc gat ccg aga ctc gaa ttc cag att      776
Gly Phe Pro Asp Phe Glu Val Thr Asp Pro Arg Leu Glu Phe Gln Ile
            225                 230                 235

ggt gtg agg gga aan nnn nnn nnn cac gtc aaa gag tac ttg gaa gaa      824
Gly Val Arg Gly Xaa Xaa Xaa Xaa His Val Lys Glu Tyr Leu Glu Glu
        240                 245                 250

atg cat cgc gaa gtt ctc tgt cat tat ccc gaa gca ttt aca gtg gga      872
Met His Arg Glu Val Leu Cys His Tyr Pro Glu Ala Phe Thr Val Gly
    255                 260                 265

gaa acc cca gga ctc aag aga ccc gag agt gca ttg ggc ctg gtc caa      920
Glu Thr Pro Gly Leu Lys Arg Pro Glu Ser Ala Leu Gly Leu Val Gln
270                 275                 280                 285

gga gga aaa cct ctt caa atg atc ttc caa ttc gag cat atg tac ttc      968
Gly Gly Lys Pro Leu Gln Met Ile Phe Gln Phe Glu His Met Tyr Phe
                290                 295                 300

gac atc caa ccc gga atg aaa ccc ttc ttc cct cgt cct tca tgg aat     1016
Asp Ile Gln Pro Gly Met Lys Pro Phe Phe Pro Arg Pro Ser Trp Asn
            305                 310                 315

ctt acc gag atc aag aaa aat ctc ggg gag tgg atg tct ttc aat gcg     1064
Leu Thr Glu Ile Lys Lys Asn Leu Gly Glu Trp Met Ser Phe Asn Ala
        320                 325                 330

gcg cat gac ggc tgg gac tct ctc tac cta gaa aat cac gat caa ccc     1112
Ala His Asp Gly Trp Asp Ser Leu Tyr Leu Glu Asn His Asp Gln Pro
    335                 340                 345

cgt att ctt ggg cgc tgg gca aaa gat act ccg gaa tgg aga gtg caa     1160
Arg Ile Leu Gly Arg Trp Ala Lys Asp Thr Pro Glu Trp Arg Val Gln
350                 355                 360                 365

gct gca aaa atg cta gcg tta ttt cat gca acc ggc aga gga acc cta     1208
Ala Ala Lys Met Leu Ala Leu Phe His Ala Thr Gly Arg Gly Thr Leu
                370                 375                 380

ttt gtc tat caa gga caa gaa atc gga acg gca aac agc aag tgg tgg     1256
Phe Val Tyr Gln Gly Gln Glu Ile Gly Thr Ala Asn Ser Lys Trp Trp
            385                 390                 395

acc aat gaa gag ttc aga gat gta gaa gag atc aat ttc ttt gcg gcc     1304
Thr Asn Glu Glu Phe Arg Asp Val Glu Glu Ile Asn Phe Phe Ala Ala
        400                 405                 410

gag aag gaa cgt cgg gag aaa gaa tcg aat ggg aaa gaa gta gac atg     1352
Glu Lys Glu Arg Arg Glu Lys Glu Ser Asn Gly Lys Glu Val Asp Met
    415                 420                 425

act gat atc tcc cga gga atc caa ggc tac gga aga gac aat tcg aga     1400
Thr Asp Ile Ser Arg Gly Ile Gln Gly Tyr Gly Arg Asp Asn Ser Arg
430                 435                 440                 445

atg ggg atg caa tgg gac gac tct cca aat ggc ggt ttc aca caa ggt     1448
Met Gly Met Gln Trp Asp Asp Ser Pro Asn Gly Gly Phe Thr Gln Gly
```

```
                450                 455                 460

aaa ccg tgg atc aag acg aat gag gag tac aag gaa atc aac gtc gca     1496
Lys Pro Trp Ile Lys Thr Asn Glu Glu Tyr Lys Glu Ile Asn Val Ala
            465                 470                 475

gcg caa gat gga gtc aag gga agc aca tta gag ttc tgg aaa cag ata     1544
Ala Gln Asp Gly Val Lys Gly Ser Thr Leu Glu Phe Trp Lys Gln Ile
        480                 485                 490

att aaa ctc cgg aaa gag aat ccg gtg ttg tgt aaa ggg gga ttt gaa     1592
Ile Lys Leu Arg Lys Glu Asn Pro Val Leu Cys Lys Gly Gly Phe Glu
    495                 500                 505

atg gtg gat cag gag aac gag gag gtt tat gcg tat gtg aga aag ggc     1640
Met Val Asp Gln Glu Asn Glu Glu Val Tyr Ala Tyr Val Arg Lys Gly
510                 515                 520                 525

gag gag aag gag tat ttg att gct tgc aat ttc aag gag cgg gat gtg     1688
Glu Glu Lys Glu Tyr Leu Ile Ala Cys Asn Phe Lys Glu Arg Asp Val
                530                 535                 540

aaa tgg aag att ccg gtt gag act ggg gag ctt ttg ttt gga agc tat     1736
Lys Trp Lys Ile Pro Val Glu Thr Gly Glu Leu Leu Phe Gly Ser Tyr
            545                 550                 555

gag gac gga gta cag ggt gga gag aag gag gga gaa tta aac ttg aga     1784
Glu Asp Gly Val Gln Gly Gly Glu Lys Glu Gly Glu Leu Asn Leu Arg
        560                 565                 570

ccg ttc gag gga aga ata tat gtt aga gat att tga                    1820
Pro Phe Glu Gly Arg Ile Tyr Val Arg Asp Ile  *
    575                 580


<210> 525
<211> 92
<212> PRT
<213> Fungi

<400> 525
Met Gly Ser Ile Ala Arg Arg Thr Ile Gly Ser Thr Pro Gln Trp Trp
 1               5                  10                  15
Lys Glu Ala Val Val Tyr Gln Ile Tyr Pro Ala Ser Tyr Leu Asp Thr
            20                  25                  30
Thr Gly Ser Gly Asp Gly Asp Leu Asn Gly Ile Thr Ser Lys Leu Pro
        35                  40                  45
Tyr Ile Arg Ser Leu Gly Val Asp Val Val Trp Ile Ser Pro Ile Tyr
    50                  55                  60
Ala Ser Pro Met Asn Asp Met Gly Tyr Asp Ile Ser Asp Tyr Arg Ala
65                  70                  75                  80
Ile Asn Pro Met Phe Gly Thr Met Glu Asp Trp Glu
                85                  90

<210> 526
<211> 492
<212> PRT
<213> Fungi

<220>
<221> UNSURE
<222> (150)...(154)
<223> Xaa = any amino acid

<400> 526
Arg Leu Cys Ala Arg Ala His Glu Leu Gly Leu Lys Leu Val Met Asp
 1               5                  10                  15
Leu Val Val Asn His Thr Ser Ser Glu His Pro Trp Phe Lys Glu Ser
```

```
            20                  25                  30
Val Ser Gly Gly Pro Asn Gly Pro Lys Arg Asp Phe Tyr Tyr Trp Gln
        35                  40                  45
Pro Pro Lys Asn Gly Lys Glu Pro Asn Asn Trp Gly Ala Met Phe Gly
    50                  55                  60
Gly Ser Ser Trp Glu Lys Asp Pro Ser His Gln Thr Asp Glu Tyr Tyr
65                  70                  75                  80
Leu His Val Tyr Asp Val Ser Gln Pro Asp Leu Asn Trp Thr Asn Pro
                85                  90                  95
Ala Val Arg Asn Glu Val Trp Asp Ile Met Arg Phe Trp Leu Asp Lys
            100                 105                 110
Gly Cys Asp Gly Phe Arg Met Asp Val Ile Asn Cys Ile Ser Lys Glu
        115                 120                 125
Pro Gly Phe Pro Asp Phe Glu Val Thr Asp Pro Arg Leu Glu Phe Gln
    130                 135                 140
Ile Gly Val Arg Gly Xaa Xaa Xaa Xaa His Val Lys Glu Tyr Leu Glu
145                 150                 155                 160
Glu Met His Arg Glu Val Leu Cys His Tyr Pro Glu Ala Phe Thr Val
                165                 170                 175
Gly Glu Thr Pro Gly Leu Lys Arg Pro Glu Ser Ala Leu Gly Leu Val
            180                 185                 190
Gln Gly Gly Lys Pro Leu Gln Met Ile Phe Gln Phe Glu His Met Tyr
        195                 200                 205
Phe Asp Ile Gln Pro Gly Met Lys Pro Phe Phe Pro Arg Pro Ser Trp
    210                 215                 220
Asn Leu Thr Glu Ile Lys Lys Asn Leu Gly Glu Trp Met Ser Phe Asn
225                 230                 235                 240
Ala Ala His Asp Gly Trp Asp Ser Leu Tyr Leu Glu Asn His Asp Gln
                245                 250                 255
Pro Arg Ile Leu Gly Arg Trp Ala Lys Asp Thr Pro Glu Trp Arg Val
            260                 265                 270
Gln Ala Ala Lys Met Leu Ala Leu Phe His Ala Thr Gly Arg Gly Thr
        275                 280                 285
Leu Phe Val Tyr Gln Gly Gln Glu Ile Gly Thr Ala Asn Ser Lys Trp
    290                 295                 300
Trp Thr Asn Glu Glu Phe Arg Asp Val Glu Glu Ile Asn Phe Phe Ala
305                 310                 315                 320
Ala Glu Lys Glu Arg Arg Glu Lys Glu Ser Asn Gly Lys Glu Val Asp
                325                 330                 335
Met Thr Asp Ile Ser Arg Gly Ile Gln Gly Tyr Gly Arg Asp Asn Ser
            340                 345                 350
Arg Met Gly Met Gln Trp Asp Asp Ser Pro Asn Gly Gly Phe Thr Gln
        355                 360                 365
Gly Lys Pro Trp Ile Lys Thr Asn Glu Glu Tyr Lys Glu Ile Asn Val
    370                 375                 380
Ala Ala Gln Asp Gly Val Lys Gly Ser Thr Leu Glu Phe Trp Lys Gln
385                 390                 395                 400
Ile Ile Lys Leu Arg Lys Glu Asn Pro Val Leu Cys Lys Gly Gly Phe
                405                 410                 415
Glu Met Val Asp Gln Glu Asn Glu Glu Val Tyr Ala Tyr Val Arg Lys
            420                 425                 430
Gly Glu Glu Lys Glu Tyr Leu Ile Ala Cys Asn Phe Lys Glu Arg Asp
        435                 440                 445
Val Lys Trp Lys Ile Pro Val Glu Thr Gly Glu Leu Leu Phe Gly Ser
    450                 455                 460
Tyr Glu Asp Gly Val Gln Gly Gly Glu Lys Glu Gly Glu Leu Asn Leu
465                 470                 475                 480
Arg Pro Phe Glu Gly Arg Ile Tyr Val Arg Asp Ile
                485                 490
```

<210> 527
<211> 1755
<212> DNA
<213> Fungi

<220>
<221> misc_feature
<222> (726)...(735)

<223> n = a, g, c or t

<400> 527
atgggatcaa tagccagaag aaccattggt tcaactcctc aatggtggaa agaagcagtc 60
gtatatcaaa tctatcccgc ttcgtatctc gacaccaccg gatctggaga tggtgatctc 120
aatggtatca cctcgaaatt gccatatatc agatctctcg gtgtagatgt agtatggatt 180
tcacctatct atgcatcccc tatgaatgat atgggctatg atatttcgga ttatcgagct 240
atcaacccta tgtttggtac tatggaagat tgggaacgtc tttgcgccag agcacacgaa 300
ctaggtttga aactcgtgat ggatctcgtg gtaaatcaca ctagttctga acacccctgg 360
ttcaaagaat ccgtctccgg cggtcccaac ggtcctaaaa gagacttcta ctactggcaa 420
ccgcccaaaa atggtaaaga gcccaataac tggggtgcca tgtttggagg ttcatcctgg 480
gaaaaagatc cttcgcatca aacagacgaa tactatctcc acgtctacga cgtcagccag 540
cccgatctta attggacaaa tccagccgtt cgcaacgagg tctgggatat catgcgcttt 600
tggcttgata agggatgtga tgggtttcgt atggatgtca tcaattgcat ttcaaaagaa 660
cctggatttc ccgattttga agtcaccgat ccgagactcg aattccagat tggtgtgagg 720
ggaaannnnn nnnnncacgt caaagagtac ttggaagaaa tgcatcgcga agttctctgt 780
cattatcccg aagcatttac agtgggagaa accccaggac tcaagagacc cgagagtgca 840
ttgggcctgg tccaaggagg aaaacctctt caaatgatct tccaattcga gcatatgtac 900
ttcgacatcc aacccggaat gaaacccttc ttccctcgtc cttcatggaa tcttaccgag 960
atcaagaaaa atctcgggga gtggatgtct ttcaatgcgg cgcatgacgg ctgggactct 1020
ctctacctag aaaatcacga tcaaccccgt attcttgggc gctgggcaaa agatactccg 1080
gaatggagag tgcaagctgc aaaaatgcta gcgttatttc atgcaaccgg cagaggaacc 1140
ctatttgtct atcaaggaca agaaatcgga acggcaaaca gcaagtggtg gaccaatgaa 1200
gagttcagag atgtagaaga gatcaatttc tttgcggccg agaaggaacg tcgggagaaa 1260
gaatcgaatg ggaaagaagt agacatgact gatatctccc gaggaatcca aggctacgga 1320
agagacaatt cgagaatggg gatgcaatgg gacgactctc caaatggcgg tttcacacaa 1380
ggtaaaccgt ggatcaagac gaatgaggag tacaaggaaa tcaacgtcgc agcgcaagat 1440
ggagtcaagg gaagcacatt agagttctgg aaacagataa ttaaactccg gaaagagaat 1500
ccggtgttgt gtaaaggggg atttgaaatg gtggatcagg agaacgagga ggtttatgcg 1560
tatgtgagaa agggcgagga gaaggagtat ttgattgctt gcaatttcaa ggagcgggat 1620
gtgaaatgga agattccggt tgagactggg gagcttttgt ttggaagcta tgaggacgga 1680
gtacagggtg gagagaagga gggagaatta aacttgagac cgttcgaggg aagaatatat 1740
gttagagata tttga 1755

<210> 528
<211> 584
<212> PRT
<213> Fungi

<220>
<221> DOMAIN
<222> (23)...(445)
<223> Alpha amylase, catalytic domain

<221> UNSURE
<222> (242)...(245)
<223> Xaa = any amino acid

<400> 528
Met Gly Ser Ile Ala Arg Arg Thr Ile Gly Ser Thr Pro Gln Trp Trp
 1   5   10   15
Lys Glu Ala Val Val Tyr Gln Ile Tyr Pro Ala Ser Tyr Leu Asp Thr
            20   25   30
Thr Gly Ser Gly Asp Gly Asp Leu Asn Gly Ile Thr Ser Lys Leu Pro
        35   40   45
Tyr Ile Arg Ser Leu Gly Val Asp Val Val Trp Ile Ser Pro Ile Tyr
    50   55   60
Ala Ser Pro Met Asn Asp Met Gly Tyr Asp Ile Ser Asp Tyr Arg Ala
65   70   75   80
Ile Asn Pro Met Phe Gly Thr Met Glu Asp Trp Glu Arg Leu Cys Ala
                85   90   95
Arg Ala His Glu Leu Gly Leu Lys Leu Val Met Asp Leu Val Val Asn
            100   105   110
His Thr Ser Ser Glu His Pro Trp Phe Lys Glu Ser Val Ser Gly Gly
        115   120   125
Pro Asn Gly Pro Lys Arg Asp Phe Tyr Tyr Trp Gln Pro Pro Lys Asn
    130   135   140
Gly Lys Glu Pro Asn Asn Trp Gly Ala Met Phe Gly Gly Ser Ser Trp

145 150 155 160
Glu Lys Asp Pro Ser His Gln Thr Asp Glu Tyr Tyr Leu His Val Tyr
165 170 175
Asp Val Ser Gln Pro Asp Leu Asn Trp Thr Asn Pro Ala Val Arg Asn
180 185 190
Glu Val Trp Asp Ile Met Arg Phe Trp Leu Asp Lys Gly Cys Asp Gly
195 200 205
Phe Arg Met Asp Val Ile Asn Cys Ile Ser Lys Glu Pro Gly Phe Pro
210 215 220
Asp Phe Glu Val Thr Asp Pro Arg Leu Glu Phe Gln Ile Gly Val Arg
225 230 235 240
Gly Xaa Xaa Xaa Xaa His Val Lys Glu Tyr Leu Glu Glu Met His Arg
245 250 255
Glu Val Leu Cys His Tyr Pro Glu Ala Phe Thr Val Gly Glu Thr Pro
260 265 270
Gly Leu Lys Arg Pro Glu Ser Ala Leu Gly Leu Val Gln Gly Gly Lys
275 280 285
Pro Leu Gln Met Ile Phe Gln Phe Glu His Met Tyr Phe Asp Ile Gln
290 295 300
Pro Gly Met Lys Pro Phe Phe Pro Arg Pro Ser Trp Asn Leu Thr Glu
305 310 315 320
Ile Lys Lys Asn Leu Gly Glu Trp Met Ser Phe Asn Ala Ala His Asp
325 330 335
Gly Trp Asp Ser Leu Tyr Leu Glu Asn His Asp Gln Pro Arg Ile Leu
340 345 350
Gly Arg Trp Ala Lys Asp Thr Pro Glu Trp Arg Val Gln Ala Ala Lys
355 360 365
Met Leu Ala Leu Phe His Ala Thr Gly Arg Gly Thr Leu Phe Val Tyr
370 375 380
Gln Gly Gln Glu Ile Gly Thr Ala Asn Ser Lys Trp Trp Thr Asn Glu
385 390 395 400
Glu Phe Arg Asp Val Glu Glu Ile Asn Phe Phe Ala Ala Glu Lys Glu
405 410 415
Arg Arg Glu Lys Glu Ser Asn Gly Lys Glu Val Asp Met Thr Asp Ile
420 425 430
Ser Arg Gly Ile Gln Gly Tyr Gly Arg Asp Asn Ser Arg Met Gly Met
435 440 445
Gln Trp Asp Asp Ser Pro Asn Gly Gly Phe Thr Gln Gly Lys Pro Trp
450 455 460
Ile Lys Thr Asn Glu Glu Tyr Lys Glu Ile Asn Val Ala Ala Gln Asp
465 470 475 480
Gly Val Lys Gly Ser Thr Leu Glu Phe Trp Lys Gln Ile Ile Lys Leu
485 490 495
Arg Lys Glu Asn Pro Val Leu Cys Lys Gly Gly Phe Glu Met Val Asp
500 505 510
Gln Glu Asn Glu Glu Val Tyr Ala Tyr Val Arg Lys Gly Glu Glu Lys
515 520 525
Glu Tyr Leu Ile Ala Cys Asn Phe Lys Glu Arg Asp Val Lys Trp Lys
530 535 540
Ile Pro Val Glu Thr Gly Glu Leu Leu Phe Gly Ser Tyr Glu Asp Gly
545 550 555 560
Val Gln Gly Gly Glu Lys Glu Gly Glu Leu Asn Leu Arg Pro Phe Glu
565 570 575
Gly Arg Ile Tyr Val Arg Asp Ile
580

<210> 529
<211> 1883
<212> DNA
<213> Cochliobolus heterostrophus ATCC 48331

<220>
<221> CDS
<222> (1)...(1265)
<223> Exon

<221> CDS
<222> (1352)...(1883)

<223> Exon

<400> 529

```
atg tca gcc gtt gca cag aag aaa gat ttg tat gtc ggc aca gac aag       48
Met Ser Ala Val Ala Gln Lys Lys Asp Leu Tyr Val Gly Thr Asp Lys
 1               5                  10                  15

cac cga tgg tgg aag gag gcc gtg gta tac cag atc tac cct gcc tca       96
His Arg Trp Trp Lys Glu Ala Val Val Tyr Gln Ile Tyr Pro Ala Ser
            20                  25                  30

ttc ctt gac agt aac ggc gat ggt tgg ggc gat gta aag ggc atc aca      144
Phe Leu Asp Ser Asn Gly Asp Gly Trp Gly Asp Val Lys Gly Ile Thr
        35                  40                  45

tcg aag ctg gat tac ctc aag gat cta ggc atc gat gtg atc tgg cta      192
Ser Lys Leu Asp Tyr Leu Lys Asp Leu Gly Ile Asp Val Ile Trp Leu
    50                  55                  60

tcc cca atc tac aaa agc ccg caa gcc gat atg ggc tac gac att gca      240
Ser Pro Ile Tyr Lys Ser Pro Gln Ala Asp Met Gly Tyr Asp Ile Ala
65                  70                  75                  80

gac tac gaa gat ata gac ccc agc tat ggt act ttg gct gat gtc gac      288
Asp Tyr Glu Asp Ile Asp Pro Ser Tyr Gly Thr Leu Ala Asp Val Asp
                85                  90                  95

aac ctg atc cac gaa acc aag aaa cgc ggc atg aag ctt gtc atg gat      336
Asn Leu Ile His Glu Thr Lys Lys Arg Gly Met Lys Leu Val Met Asp
            100                 105                 110

ctg gtt gtc aac cac acc tca gaa gag cat gca tgg ttt ctc gat tct      384
Leu Val Val Asn His Thr Ser Glu Glu His Ala Trp Phe Leu Asp Ser
        115                 120                 125

cga tca tcc aag gaa tca aag aaa aga gac tgg tat atc tgg aag cca      432
Arg Ser Ser Lys Glu Ser Lys Lys Arg Asp Trp Tyr Ile Trp Lys Pro
    130                 135                 140

gca aag tac gac gcc gat ggc aac cgt cag ccg cca aac aac tgg gca      480
Ala Lys Tyr Asp Ala Asp Gly Asn Arg Gln Pro Pro Asn Asn Trp Ala
145                 150                 155                 160

cag atc ttg gga gaa gcc aac tcg gcg tgg aca tgg gat gaa aag aca      528
Gln Ile Leu Gly Glu Ala Asn Ser Ala Trp Thr Trp Asp Glu Lys Thr
                165                 170                 175

cag gag tat tac ctt tca ctc ttc aca ccg gag cag ccc gac ctc aac      576
Gln Glu Tyr Tyr Leu Ser Leu Phe Thr Pro Glu Gln Pro Asp Leu Asn
            180                 185                 190

tgg gaa aat cct gat gtc cgt gag gcc gta cat aac atc ctc cgc ttc      624
Trp Glu Asn Pro Asp Val Arg Glu Ala Val His Asn Ile Leu Arg Phe
        195                 200                 205

tgg cta gat cgt ggc gtc tcc ggt ttc cgc atg gat gtc atc aac ctt      672
Trp Leu Asp Arg Gly Val Ser Gly Phe Arg Met Asp Val Ile Asn Leu
    210                 215                 220

att tca aag gtc caa acc ttc ccc gat gca cca atc tca gtc aaa gat      720
Ile Ser Lys Val Gln Thr Phe Pro Asp Ala Pro Ile Ser Val Lys Asp
225                 230                 235                 240

gcc aag tac caa cca ggc gac aag tac ttt gcc aac ggc ccc cgc cta      768
Ala Lys Tyr Gln Pro Gly Asp Lys Tyr Phe Ala Asn Gly Pro Arg Leu
                245                 250                 255

cac gaa tgg ctc aag gaa ctg cgc cgc gat gtc ctg tca aag tat gac      816
```

```
His Glu Trp Leu Lys Glu Leu Arg Arg Asp Val Leu Ser Lys Tyr Asp
            260                 265                 270

act ctg acc gta ggc gaa atg ccc ttt gtc cgc gat gaa gac gaa gtc      864
Thr Leu Thr Val Gly Glu Met Pro Phe Val Arg Asp Glu Asp Glu Val
        275                 280                 285

atc aaa gtc gtc ggc tca gaa agc ggt gag ctg aac atg att ttc aac      912
Ile Lys Val Val Gly Ser Glu Ser Gly Glu Leu Asn Met Ile Phe Asn
    290                 295                 300

ttc gac ctt gtc gac att gac aac gtt cct ggc gac ttc aag tac aca      960
Phe Asp Leu Val Asp Ile Asp Asn Val Pro Gly Asp Phe Lys Tyr Thr
305                 310                 315                 320

ctg cac ccc tgg gac gca cgc gat ctc aag aaa atc gta aac aga cta     1008
Leu His Pro Trp Asp Ala Arg Asp Leu Lys Lys Ile Val Asn Arg Leu
                325                 330                 335

cag cgt ctg atg ctt gag cgt gac ggt tgg aac aca cta tat gtg gaa     1056
Gln Arg Leu Met Leu Glu Arg Asp Gly Trp Asn Thr Leu Tyr Val Glu
            340                 345                 350

aac cac gac cag ccc cgc agc ata agc aga tat act gac gac agc gat     1104
Asn His Asp Gln Pro Arg Ser Ile Ser Arg Tyr Thr Asp Asp Ser Asp
        355                 360                 365

gag tgg cgc tcc tac ggc gca aag ctg ctc tgc ctt atg cag aca acc     1152
Glu Trp Arg Ser Tyr Gly Ala Lys Leu Leu Cys Leu Met Gln Thr Thr
    370                 375                 380

ctc gcc gga acc ctc tac gtc tat caa ggc gag gaa att ggc atg cgc     1200
Leu Ala Gly Thr Leu Tyr Val Tyr Gln Gly Glu Glu Ile Gly Met Arg
385                 390                 395                 400

aac ata ccc aaa gaa tgg ggc cca gaa gag tac aaa gac att gaa agt     1248
Asn Ile Pro Lys Glu Trp Gly Pro Glu Glu Tyr Lys Asp Ile Glu Ser
                405                 410                 415

atc aat ttc ttc aaa aa gtacgtgtcc ccatcttctc ccctaacctc             1295
Ile Asn Phe Phe Lys Lys
            420

cttccatccc tccccctttt gtatctcgaa acgacgcttt gaactaaacg acccag a     1352
tac cgc gac ctt tac ccc aac aac ccc gaa aaa caa gcc ctc gca gca     1400
Tyr Arg Asp Leu Tyr Pro Asn Asn Pro Glu Lys Gln Ala Leu Ala Ala
        425                 430                 435

gaa atc atg caa cgc aaa gcc cgc gac agc gcc cga aca ccc atg caa     1448
Glu Ile Met Gln Arg Lys Ala Arg Asp Ser Ala Arg Thr Pro Met Gln
    440                 445                 450

tgg aac gcc gac tcc caa gca ggc ttc aca tcc gga acc ccc tgg atg     1496
Trp Asn Ala Asp Ser Gln Ala Gly Phe Thr Ser Gly Thr Pro Trp Met
455                 460                 465                 470

cgc gtc aac acg gac tat ccc gcc atc aac gtc gcc gag caa ctc gcc     1544
Arg Val Asn Thr Asp Tyr Pro Ala Ile Asn Val Ala Glu Gln Leu Ala
                475                 480                 485

aac ccc acg ccc gcc ccc ggt act ctg tcc gtc cat gcg ttc tgg aag     1592
Asn Pro Thr Pro Ala Pro Gly Thr Leu Ser Val His Ala Phe Trp Lys
            490                 495                 500

cgt gcc ctg gag tac cgc aag gaa aac aag gac gtt ttc gtg tac ggg     1640
Arg Ala Leu Glu Tyr Arg Lys Glu Asn Lys Asp Val Phe Val Tyr Gly
        505                 510                 515
```

```
gat ttt gaa atg ctg gat atg gga cat gag aag gtt gtg gcg ttt agg     1688
Asp Phe Glu Met Leu Asp Met Gly His Glu Lys Val Val Ala Phe Arg
    520                 525                 530

cgg tgg ggc aag gat aag gcg ttt atc act gtg ctg aat ttt agc ggt     1736
Arg Trp Gly Lys Asp Lys Ala Phe Ile Thr Val Leu Asn Phe Ser Gly
535                 540                 545                 550

gag gag gtt aag tgg gga ggt ttg ggt gaa ata aag gtg aag aag tgg     1784
Glu Glu Val Lys Trp Gly Gly Leu Gly Glu Ile Lys Val Lys Lys Trp
                555                 560                 565

gtt gcg ggg aat tat gat gag agg gag ttg gaa acg aag gcc aag agt     1832
Val Ala Gly Asn Tyr Asp Glu Arg Glu Leu Glu Thr Lys Ala Lys Ser
            570                 575                 580

ggg gcg ttg gag ttg agg gct tgg gag ggg ttg ctg gga ctt ttg gag     1880
Gly Ala Leu Glu Leu Arg Ala Trp Glu Gly Leu Leu Gly Leu Leu Glu
        585                 590                 595

tga                                                                 1883
 *
```

<210> 530
<211> 422
<212> PRT
<213> Cochliobolus heterostrophus ATCC 48331

<400> 530

```
Met Ser Ala Val Ala Gln Lys Lys Asp Leu Tyr Val Gly Thr Asp Lys
 1               5                  10                  15
His Arg Trp Trp Lys Glu Ala Val Val Tyr Gln Ile Tyr Pro Ala Ser
            20                  25                  30
Phe Leu Asp Ser Asn Gly Asp Gly Trp Gly Asp Val Lys Gly Ile Thr
        35                  40                  45
Ser Lys Leu Asp Tyr Leu Lys Asp Leu Gly Ile Asp Val Ile Trp Leu
    50                  55                  60
Ser Pro Ile Tyr Lys Ser Pro Gln Ala Asp Met Gly Tyr Asp Ile Ala
65                  70                  75                  80
Asp Tyr Glu Asp Ile Asp Pro Ser Tyr Gly Thr Leu Ala Asp Val Asp
                85                  90                  95
Asn Leu Ile His Glu Thr Lys Lys Arg Gly Met Lys Leu Val Met Asp
            100                 105                 110
Leu Val Val Asn His Thr Ser Glu Glu His Ala Trp Phe Leu Asp Ser
        115                 120                 125
Arg Ser Ser Lys Glu Ser Lys Lys Arg Asp Trp Tyr Ile Trp Lys Pro
    130                 135                 140
Ala Lys Tyr Asp Ala Asp Gly Asn Arg Gln Pro Pro Asn Asn Trp Ala
145                 150                 155                 160
Gln Ile Leu Gly Glu Ala Asn Ser Ala Trp Thr Trp Asp Glu Lys Thr
                165                 170                 175
Gln Glu Tyr Tyr Leu Ser Leu Phe Thr Pro Glu Gln Pro Asp Leu Asn
            180                 185                 190
Trp Glu Asn Pro Asp Val Arg Glu Ala Val His Asn Ile Leu Arg Phe
        195                 200                 205
Trp Leu Asp Arg Gly Val Ser Gly Phe Arg Met Asp Val Ile Asn Leu
    210                 215                 220
Ile Ser Lys Val Gln Thr Phe Pro Asp Ala Pro Ile Ser Val Lys Asp
225                 230                 235                 240
Ala Lys Tyr Gln Pro Gly Asp Lys Tyr Phe Ala Asn Gly Pro Arg Leu
                245                 250                 255
His Glu Trp Leu Lys Glu Leu Arg Arg Asp Val Leu Ser Lys Tyr Asp
            260                 265                 270
Thr Leu Thr Val Gly Glu Met Pro Phe Val Arg Asp Glu Asp Glu Val
        275                 280                 285
Ile Lys Val Val Gly Ser Glu Ser Gly Glu Leu Asn Met Ile Phe Asn
```

290 295 300
Phe Asp Leu Val Asp Ile Asp Asn Val Pro Gly Asp Phe Lys Tyr Thr
305 310 315 320
Leu His Pro Trp Asp Ala Arg Asp Leu Lys Lys Ile Val Asn Arg Leu
325 330 335
Gln Arg Leu Met Leu Glu Arg Asp Gly Trp Asn Thr Leu Tyr Val Glu
340 345 350
Asn His Asp Gln Pro Arg Ser Ile Ser Arg Tyr Thr Asp Asp Ser Asp
355 360 365
Glu Trp Arg Ser Tyr Gly Ala Lys Leu Leu Cys Leu Met Gln Thr Thr
370 375 380
Leu Ala Gly Thr Leu Tyr Val Tyr Gln Gly Glu Glu Ile Gly Met Arg
385 390 395 400
Asn Ile Pro Lys Glu Trp Gly Pro Glu Glu Tyr Lys Asp Ile Glu Ser
405 410 415
Ile Asn Phe Phe Lys Lys
420

<210> 531
<211> 176
<212> PRT
<213> Cochliobolus heterostrophus ATCC 48331

<400> 531
Tyr Arg Asp Leu Tyr Pro Asn Asn Pro Glu Lys Gln Ala Leu Ala Ala
1 5 10 15
Glu Ile Met Gln Arg Lys Ala Arg Asp Ser Ala Arg Thr Pro Met Gln
20 25 30
Trp Asn Ala Asp Ser Gln Ala Gly Phe Thr Ser Gly Thr Pro Trp Met
35 40 45
Arg Val Asn Thr Asp Tyr Pro Ala Ile Asn Val Ala Glu Gln Leu Ala
50 55 60
Asn Pro Thr Pro Ala Pro Gly Thr Leu Ser Val His Ala Phe Trp Lys
65 70 75 80
Arg Ala Leu Glu Tyr Arg Lys Glu Asn Lys Asp Val Phe Val Tyr Gly
85 90 95
Asp Phe Glu Met Leu Asp Met Gly His Glu Lys Val Val Ala Phe Arg
100 105 110
Arg Trp Gly Lys Asp Lys Ala Phe Ile Thr Val Leu Asn Phe Ser Gly
115 120 125
Glu Glu Val Lys Trp Gly Gly Leu Gly Glu Ile Lys Val Lys Lys Trp
130 135 140
Val Ala Gly Asn Tyr Asp Glu Arg Glu Leu Glu Thr Lys Ala Lys Ser
145 150 155 160
Gly Ala Leu Glu Leu Arg Ala Trp Glu Gly Leu Leu Gly Leu Leu Glu
165 170 175

<210> 532
<211> 1797
<212> DNA
<213> Cochliobolus heterostrophus ATCC 48331

<400> 532
atgtcagccg ttgcacagaa gaaagatttg tatgtcggca cagacaagca ccgatggtgg 60
aaggaggccg tggtatacca gatctaccct gcctcattcc ttgacagtaa cggcgatggt 120
tggggcgatg taaagggcat cacatcgaag ctggattacc tcaaggatct aggcatcgat 180
gtgatctggc tatccccaat ctacaaaagc ccgcaagccg atatgggcta cgacattgca 240
gactacgaag atatagaccc cagctatggt actttggctg atgtcgacaa cctgatccac 300
gaaaccaaga aacgcggcat gaagcttgtc atggatctgg ttgtcaacca cacctcagaa 360
gagcatgcat ggtttctcga ttctcgatca tccaaggaat caaagaaaag agactggtat 420
atctggaagc cagcaaagta cgacgccgat ggcaaccgtc agccgccaaa caactgggca 480
cagatcttgg gagaagccaa ctcggcgtgg acatgggatg aaaagacaca ggagtattac 540
ctttcactct tcacaccgga gcagcccgac ctcaactggg aaaatcctga tgtccgtgag 600
gccgtacata acatcctccg cttctggcta gatcgtggcg tctccggttt ccgcatggat 660
gtcatcaacc ttatttcaaa ggtccaaacc ttccccgatg caccaatctc agtcaaagat 720
gccaagtacc aaccaggcga caagtacttt gccaacggcc cccgcctaca cgaatggctc 780
aaggaactgc gccgcgatgt cctgtcaaag tatgacactc tgaccgtagg cgaaatgccc 840
tttgtccgcg atgaagacga agtcatcaaa gtcgtcggct cagaaagcgg tgagctgaac 900

atgattttca acttcgacct tgtcgacatt gacaacgttc ctggcgactt caagtacaca 960
ctgcacccct gggacgcacg cgatctcaag aaaatcgtaa acagactaca gcgtctgatg 1020
cttgagcgtg acggttggaa cacactatat gtggaaaacc acgaccagcc ccgcagcata 1080
agcagatata ctgacgacag cgatgagtgg cgctcctacg gcgcaaagct gctctgcctt 1140
atgcagacaa ccctcgccgg aaccctctac gtctatcaag gcgaggaaat tggcatgcgc 1200
aacataccca aagaatgggg cccagaagag tacaaagaca ttgaaagtat caatttcttc 1260
aaaaaatacc gcgaccttta ccccaacaac cccgaaaaac aagccctcgc agcagaaatc 1320
atgcaacgca aagcccgcga cagcgcccga acacccatgc aatggaacgc cgactcccaa 1380
gcaggcttca catccggaac cccctggatg cgcgtcaaca cggactatcc cgccatcaac 1440
gtcgccgagc aactcgccaa ccccacgccc gcccccggta ctctgtccgt ccatgcgttc 1500
tggaagcgtg ccctggagta ccgcaaggaa aacaaggacg ttttcgtgta cggggatttt 1560
gaaatgctgg atatgggaca tgagaaggtt gtggcgttta ggcggtgggg caaggataag 1620
gcgtttatca ctgtgctgaa ttttagcggt gaggaggtta agtggggagg tttgggtgaa 1680
ataaaggtga agaagtgggt tgcggggaat tatgatgaga gggagttgga aacgaaggcc 1740
aagagtgggg cgttggagtt gagggcttgg gaggggttgc tgggactttt ggagtga 1797

<210> 533
<211> 598
<212> PRT
<213> Cochliobolus heterostrophus ATCC 48331

<220>
<221> DOMAIN
<222> (27)...(450)
<223> Alpha amylase, catalytic domain

<400> 533
Met Ser Ala Val Ala Gln Lys Lys Asp Leu Tyr Val Gly Thr Asp Lys
 1   5   10   15
His Arg Trp Trp Lys Glu Ala Val Val Tyr Gln Ile Tyr Pro Ala Ser
            20   25   30
Phe Leu Asp Ser Asn Gly Asp Gly Trp Gly Asp Val Lys Gly Ile Thr
        35   40   45
Ser Lys Leu Asp Tyr Leu Lys Asp Leu Gly Ile Asp Val Ile Trp Leu
    50   55   60
Ser Pro Ile Tyr Lys Ser Pro Gln Ala Asp Met Gly Tyr Asp Ile Ala
65   70   75   80
Asp Tyr Glu Asp Ile Asp Pro Ser Tyr Gly Thr Leu Ala Asp Val Asp
                85   90   95
Asn Leu Ile His Glu Thr Lys Lys Arg Gly Met Lys Leu Val Met Asp
            100   105   110
Leu Val Val Asn His Thr Ser Glu Glu His Ala Trp Phe Leu Asp Ser
        115   120   125
Arg Ser Ser Lys Glu Ser Lys Lys Arg Asp Trp Tyr Ile Trp Lys Pro
    130   135   140
Ala Lys Tyr Asp Ala Asp Gly Asn Arg Gln Pro Pro Asn Asn Trp Ala
145   150   155   160
Gln Ile Leu Gly Glu Ala Asn Ser Ala Trp Thr Trp Asp Glu Lys Thr
                165   170   175
Gln Glu Tyr Tyr Leu Ser Leu Phe Thr Pro Glu Gln Pro Asp Leu Asn
            180   185   190
Trp Glu Asn Pro Asp Val Arg Glu Ala Val His Asn Ile Leu Arg Phe
        195   200   205
Trp Leu Asp Arg Gly Val Ser Gly Phe Arg Met Asp Val Ile Asn Leu
    210   215   220
Ile Ser Lys Val Gln Thr Phe Pro Asp Ala Pro Ile Ser Val Lys Asp
225   230   235   240
Ala Lys Tyr Gln Pro Gly Asp Lys Tyr Phe Ala Asn Gly Pro Arg Leu
                245   250   255
His Glu Trp Leu Lys Glu Leu Arg Arg Asp Val Leu Ser Lys Tyr Asp
            260   265   270
Thr Leu Thr Val Gly Glu Met Pro Phe Val Arg Asp Glu Asp Glu Val
        275   280   285
Ile Lys Val Val Gly Ser Glu Ser Gly Glu Leu Asn Met Ile Phe Asn
    290   295   300
Phe Asp Leu Val Asp Ile Asp Asn Val Pro Gly Asp Phe Lys Tyr Thr
305   310   315   320
Leu His Pro Trp Asp Ala Arg Asp Leu Lys Lys Ile Val Asn Arg Leu

```
                325                 330                 335
Gln Arg Leu Met Leu Glu Arg Asp Gly Trp Asn Thr Leu Tyr Val Glu
            340                 345                 350
Asn His Asp Gln Pro Arg Ser Ile Ser Arg Tyr Thr Asp Asp Ser Asp
        355                 360                 365
Glu Trp Arg Ser Tyr Gly Ala Lys Leu Leu Cys Leu Met Gln Thr Thr
    370                 375                 380
Leu Ala Gly Thr Leu Tyr Val Tyr Gln Gly Glu Glu Ile Gly Met Arg
385                 390                 395                 400
Asn Ile Pro Lys Glu Trp Gly Pro Glu Glu Tyr Lys Asp Ile Glu Ser
                405                 410                 415
Ile Asn Phe Phe Lys Lys Tyr Arg Asp Leu Tyr Pro Asn Asn Pro Glu
            420                 425                 430
Lys Gln Ala Leu Ala Ala Glu Ile Met Gln Arg Lys Ala Arg Asp Ser
        435                 440                 445
Ala Arg Thr Pro Met Gln Trp Asn Ala Asp Ser Gln Ala Gly Phe Thr
    450                 455                 460
Ser Gly Thr Pro Trp Met Arg Val Asn Thr Asp Tyr Pro Ala Ile Asn
465                 470                 475                 480
Val Ala Glu Gln Leu Ala Asn Pro Thr Pro Ala Pro Gly Thr Leu Ser
                485                 490                 495
Val His Ala Phe Trp Lys Arg Ala Leu Glu Tyr Arg Lys Glu Asn Lys
            500                 505                 510
Asp Val Phe Val Tyr Gly Asp Phe Glu Met Leu Asp Met Gly His Glu
        515                 520                 525
Lys Val Val Ala Phe Arg Arg Trp Gly Lys Asp Lys Ala Phe Ile Thr
    530                 535                 540
Val Leu Asn Phe Ser Gly Glu Glu Val Lys Trp Gly Gly Leu Gly Glu
545                 550                 555                 560
Ile Lys Val Lys Lys Trp Val Ala Gly Asn Tyr Asp Glu Arg Glu Leu
                565                 570                 575
Glu Thr Lys Ala Lys Ser Gly Ala Leu Glu Leu Arg Ala Trp Glu Gly
            580                 585                 590
Leu Leu Gly Leu Leu Glu
        595
```

<210> 534
<211> 1922
<212> DNA
<213> Cochliobolus heterostrophus ATCC 48331

<220>
<221> CDS
<222> (1)...(262)
<223> Exon

<221> CDS
<222> (316)...(425)
<223> Exon

<221> CDS
<222> (483)...(1002)
<223> Exon

<221> CDS
<222> (1045)...(1922)
<223> Exon

<400> 534

```
atg ttg ttg ctc aac atc ttc acc acc ctc ttc ttc tac atc acc tgc       48
Met Leu Leu Leu Asn Ile Phe Thr Thr Leu Phe Phe Tyr Ile Thr Cys
 1               5                  10                  15

atc gtc tcc gcc gcc gac acc aat gct tgg aag tcc cgc agc atc tac       96
Ile Val Ser Ala Ala Asp Thr Asn Ala Trp Lys Ser Arg Ser Ile Tyr
             20                  25                  30

ttt gtc ctg acg gat cgt att gcc cgc aac agc agc gac acg ggc ggc      144
```

```
Phe Val Leu Thr Asp Arg Ile Ala Arg Asn Ser Ser Asp Thr Gly Gly
         35                  40                  45

tca gcg tgc agc gac ctc ggc aac tac tgc ggt gga act ttc cag ggc      192
Ser Ala Cys Ser Asp Leu Gly Asn Tyr Cys Gly Gly Thr Phe Gln Gly
     50                  55                  60

ctc gag tct aag ctc gac tac atc aag gga ctt gga ttc gat gcc att      240
Leu Glu Ser Lys Leu Asp Tyr Ile Lys Gly Leu Gly Phe Asp Ala Ile
 65                  70                  75                  80

tgg att acc ccc gtc gtc tca a gtatgttcat cacggcttct gccaacacag       292
Trp Ile Thr Pro Val Val Ser
                 85

agcagtacta atccaacttg tag ac aag gct gct gga tac cat ggc tac tgg     344
                        Asn Lys Ala Ala Gly Tyr His Gly Tyr Trp
                                90                  95

gcc gag gac ttg tat gcc gtc aac tca aac tac ggc act gct gcc gac      392
Ala Glu Asp Leu Tyr Ala Val Asn Ser Asn Tyr Gly Thr Ala Ala Asp
        100                 105                 110

ttg aag agc ttg gtt gcc gct gcc cat gcc aag gtatgtccgg atttagtatc    445
Leu Lys Ser Leu Val Ala Ala Ala His Ala Lys
    115                 120

tatgttcata gctagtacat ttgcttactc atcatag ggc atc tac atg atg gtc     500
                                        Gly Ile Tyr Met Met Val
                                        125                 130

gac gtt gtc gca aac cac atg ggt cct gga gca atc aca aac aac cgc      548
Asp Val Val Ala Asn His Met Gly Pro Gly Ala Ile Thr Asn Asn Arg
                135                 140                 145

cct gaa cct ctc aac cag gct tca tca tac cac cct cct tgc aac atc      596
Pro Glu Pro Leu Asn Gln Ala Ser Ser Tyr His Pro Pro Cys Asn Ile
            150                 155                 160

gac tac aac aac caa acc agt gtc gag gta tgt caa ata gcc gga ctc      644
Asp Tyr Asn Asn Gln Thr Ser Val Glu Val Cys Gln Ile Ala Gly Leu
        165                 170                 175

ccc gac atc tac acc acc aag agc gag atc cgc acg ctc ctc aac acc      692
Pro Asp Ile Tyr Thr Thr Lys Ser Glu Ile Arg Thr Leu Leu Asn Thr
    180                 185                 190

tgg gtc aac tgg ctc gta aac gag tac agc ttc gac ggt gtc cgc atc      740
Trp Val Asn Trp Leu Val Asn Glu Tyr Ser Phe Asp Gly Val Arg Ile
195                 200                 205                 210

gac acc gtc aag cac gtc gaa aag gac ttt tgg cct ggc ttc tct gcc      788
Asp Thr Val Lys His Val Glu Lys Asp Phe Trp Pro Gly Phe Ser Ala
                215                 220                 225

gct acc ggt gtc tac aac att ggc gag gtg ttt gac gga gac cca gcc      836
Ala Thr Gly Val Tyr Asn Ile Gly Glu Val Phe Asp Gly Asp Pro Ala
            230                 235                 240

tac ctt gcc ccg tac gcc aag ctt atg ccc ggc ctc ctc aac tac gca      884
Tyr Leu Ala Pro Tyr Ala Lys Leu Met Pro Gly Leu Leu Asn Tyr Ala
        245                 250                 255

gtc tac tac ccg atg aac aac ttt tac cag caa acg ggc tct tcc cag      932
Val Tyr Tyr Pro Met Asn Asn Phe Tyr Gln Gln Thr Gly Ser Ser Gln
    260                 265                 270

gcg ctt gta gac atg atg aac act gtc agc aac acc ttc cct gac cca      980
```

```
Ala Leu Val Asp Met Met Asn Thr Val Ser Asn Thr Phe Pro Asp Pro
275                 280                 285                 290

tct gcc ttg gga acc ttc ctc g gtctagaact agtggatccc ccgggctgca      1032
Ser Ala Leu Gly Thr Phe Leu
                295

ggaattcgat ag ac aac cac gac aac aag cgc tgg ttg aac gtc aag aac     1082
             Asp Asn His Asp Asn Lys Arg Trp Leu Asn Val Lys Asn
                    300                 305                 310

gac cag act ctg ctc aag aac gct ctt gct tat gtc atc ctc gca cgt     1130
Asp Gln Thr Leu Leu Lys Asn Ala Leu Ala Tyr Val Ile Leu Ala Arg
                315                 320                 325

ggt atc ccc atc ttg tac tat ggt acc gag cag gga tac gct ggt ggt     1178
Gly Ile Pro Ile Leu Tyr Tyr Gly Thr Glu Gln Gly Tyr Ala Gly Gly
            330                 335                 340

gac gac cca gct aac cga gag gat ctg tgg cgc agt ggc ttc aac acc     1226
Asp Asp Pro Ala Asn Arg Glu Asp Leu Trp Arg Ser Gly Phe Asn Thr
        345                 350                 355

aat gcc aac ctc tac caa gcc atc aag aaa ctg acc gcc gcc cga cag     1274
Asn Ala Asn Leu Tyr Gln Ala Ile Lys Lys Leu Thr Ala Ala Arg Gln
    360                 365                 370

gct gcc ggt ggt ctc gca gga aac gac cac gtc cac ctg tac gtc gcc     1322
Ala Ala Gly Gly Leu Ala Gly Asn Asp His Val His Leu Tyr Val Ala
375                 380                 385                 390

gac acg gct tac gcc tgg agc cgt gcc aac ggc aac ctg att gtc ctc     1370
Asp Thr Ala Tyr Ala Trp Ser Arg Ala Asn Gly Asn Leu Ile Val Leu
                395                 400                 405

acc acc aac gct ggc ggc aac tcc aac acc cag cac tgc ttc aac acg     1418
Thr Thr Asn Ala Gly Gly Asn Ser Asn Thr Gln His Cys Phe Asn Thr
            410                 415                 420

caa aag gca aac ggc cgc tgg acc aac gtc tac ggc aac ggc gcc acc     1466
Gln Lys Ala Asn Gly Arg Trp Thr Asn Val Tyr Gly Asn Gly Ala Thr
        425                 430                 435

gtc tct gcc gat agc aac ggc caa atc tgc gtc tcc gtc aca aac ggc     1514
Val Ser Ala Asp Ser Asn Gly Gln Ile Cys Val Ser Val Thr Asn Gly
    440                 445                 450

gag ccc gtt gtc ctc ctc gcc ggc tcc gct acc ccc acc act ggc act     1562
Glu Pro Val Val Leu Leu Ala Gly Ser Ala Thr Pro Thr Thr Gly Thr
455                 460                 465                 470

acc ctc tcc acc cgc acc gcc act gcc acc gcc aca cca acc gca tgc     1610
Thr Leu Ser Thr Arg Thr Ala Thr Ala Thr Ala Thr Pro Thr Ala Cys
                475                 480                 485

ccc acc gcc gtc tcc gtc tcc ttc acc cac cgc gtc acc act gtt ccc     1658
Pro Thr Ala Val Ser Val Ser Phe Thr His Arg Val Thr Thr Val Pro
            490                 495                 500

ggt gac acc atc aaa atc act ggc aac acg gcc cag cta ggt aac tgg     1706
Gly Asp Thr Ile Lys Ile Thr Gly Asn Thr Ala Gln Leu Gly Asn Trp
        505                 510                 515

act ccc gcc aac ggc ctt gcc ttg tcc gca gct agc tac aca tcc agc     1754
Thr Pro Ala Asn Gly Leu Ala Leu Ser Ala Ala Ser Tyr Thr Ser Ser
    520                 525                 530

aac cct atc tgg acc att acc gtg ccc ctg gcc gct gga tcc tcc atc     1802
```

```
Asn Pro Ile Trp Thr Ile Thr Val Pro Leu Ala Ala Gly Ser Ser Ile
535                 540                 545                 550

tcg tac aag ttt gtc aag att gac agt gga gga act gtc acc tgg gag     1850
Ser Tyr Lys Phe Val Lys Ile Asp Ser Gly Gly Thr Val Thr Trp Glu
                555                 560                 565

agt gac ccc aac agg tca tac act gcg ccg agc tgc cag gcg agt gcc     1898
Ser Asp Pro Asn Arg Ser Tyr Thr Ala Pro Ser Cys Gln Ala Ser Ala
            570                 575                 580

agt gtg aac agc tca tgg caa tag                                    1922
Ser Val Asn Ser Ser Trp Gln  *
        585


<210> 535
<211> 87
<212> PRT
<213> Cochliobolus heterostrophus ATCC 48331

<400> 535
Met Leu Leu Leu Asn Ile Phe Thr Thr Leu Phe Phe Tyr Ile Thr Cys
 1               5                  10                  15
Ile Val Ser Ala Ala Asp Thr Asn Ala Trp Lys Ser Arg Ser Ile Tyr
            20                  25                  30
Phe Val Leu Thr Asp Arg Ile Ala Arg Asn Ser Ser Asp Thr Gly Gly
        35                  40                  45
Ser Ala Cys Ser Asp Leu Gly Asn Tyr Cys Gly Gly Thr Phe Gln Gly
    50                  55                  60
Leu Glu Ser Lys Leu Asp Tyr Ile Lys Gly Leu Gly Phe Asp Ala Ile
65                  70                  75                  80
Trp Ile Thr Pro Val Val Ser
                85

<210> 536
<211> 37
<212> PRT
<213> Cochliobolus heterostrophus ATCC 48331

<400> 536
Asn Lys Ala Ala Gly Tyr His Gly Tyr Trp Ala Glu Asp Leu Tyr Ala
 1               5                  10                  15
Val Asn Ser Asn Tyr Gly Thr Ala Ala Asp Leu Lys Ser Leu Val Ala
            20                  25                  30
Ala Ala His Ala Lys
        35

<210> 537
<211> 173
<212> PRT
<213> Cochliobolus heterostrophus ATCC 48331

<400> 537
Gly Ile Tyr Met Met Val Asp Val Val Ala Asn His Met Gly Pro Gly
 1               5                  10                  15
Ala Ile Thr Asn Asn Arg Pro Glu Pro Leu Asn Gln Ala Ser Ser Tyr
            20                  25                  30
His Pro Pro Cys Asn Ile Asp Tyr Asn Asn Gln Thr Ser Val Glu Val
        35                  40                  45
Cys Gln Ile Ala Gly Leu Pro Asp Ile Tyr Thr Thr Lys Ser Glu Ile
    50                  55                  60
Arg Thr Leu Leu Asn Thr Trp Val Asn Trp Leu Val Asn Glu Tyr Ser
65                  70                  75                  80
Phe Asp Gly Val Arg Ile Asp Thr Val Lys His Val Glu Lys Asp Phe
                85                  90                  95
Trp Pro Gly Phe Ser Ala Ala Thr Gly Val Tyr Asn Ile Gly Glu Val
            100                 105                 110
```

Phe Asp Gly Asp Pro Ala Tyr Leu Ala Pro Tyr Ala Lys Leu Met Pro
115 120 125
Gly Leu Leu Asn Tyr Ala Val Tyr Tyr Pro Met Asn Asn Phe Tyr Gln
130 135 140
Gln Thr Gly Ser Ser Gln Ala Leu Val Asp Met Met Asn Thr Val Ser
145 150 155 160
Asn Thr Phe Pro Asp Pro Ser Ala Leu Gly Thr Phe Leu
165 170

<210> 538
<211> 292
<212> PRT
<213> Cochliobolus heterostrophus ATCC 48331

<400> 538
Asp Asn His Asp Asn Lys Arg Trp Leu Asn Val Lys Asn Asp Gln Thr
1 5 10 15
Leu Leu Lys Asn Ala Leu Ala Tyr Val Ile Leu Ala Arg Gly Ile Pro
20 25 30
Ile Leu Tyr Tyr Gly Thr Glu Gln Gly Tyr Ala Gly Gly Asp Asp Pro
35 40 45
Ala Asn Arg Glu Asp Leu Trp Arg Ser Gly Phe Asn Thr Asn Ala Asn
50 55 60
Leu Tyr Gln Ala Ile Lys Lys Leu Thr Ala Ala Arg Gln Ala Ala Gly
65 70 75 80
Gly Leu Ala Gly Asn Asp His Val His Leu Tyr Val Ala Asp Thr Ala
85 90 95
Tyr Ala Trp Ser Arg Ala Asn Gly Asn Leu Ile Val Leu Thr Thr Asn
100 105 110
Ala Gly Gly Asn Ser Asn Thr Gln His Cys Phe Asn Thr Gln Lys Ala
115 120 125
Asn Gly Arg Trp Thr Asn Val Tyr Gly Asn Gly Ala Thr Val Ser Ala
130 135 140
Asp Ser Asn Gly Gln Ile Cys Val Ser Val Thr Asn Gly Glu Pro Val
145 150 155 160
Val Leu Leu Ala Gly Ser Ala Thr Pro Thr Thr Gly Thr Thr Leu Ser
165 170 175
Thr Arg Thr Ala Thr Ala Thr Ala Thr Pro Thr Ala Cys Pro Thr Ala
180 185 190
Val Ser Val Ser Phe Thr His Arg Val Thr Thr Val Pro Gly Asp Thr
195 200 205
Ile Lys Ile Thr Gly Asn Thr Ala Gln Leu Gly Asn Trp Thr Pro Ala
210 215 220
Asn Gly Leu Ala Leu Ser Ala Ala Ser Tyr Thr Ser Ser Asn Pro Ile
225 230 235 240
Trp Thr Ile Thr Val Pro Leu Ala Ala Gly Ser Ser Ile Ser Tyr Lys
245 250 255
Phe Val Lys Ile Asp Ser Gly Gly Thr Val Thr Trp Glu Ser Asp Pro
260 265 270
Asn Arg Ser Tyr Thr Ala Pro Ser Cys Gln Ala Ser Ala Ser Val Asn
275 280 285
Ser Ser Trp Gln
290

<210> 539
<211> 1770
<212> DNA
<213> Cochliobolus heterostrophus ATCC 48331

<400> 539
atgttgttgc tcaacatctt caccaccctc ttcttctaca tcacctgcat cgtctccgcc 60
gccgacacca atgcttggaa gtcccgcagc atctactttg tcctgacgga tcgtattgcc 120
cgcaacagca gcgacacggg cggctcagcg tgcagcgacc tcggcaacta ctgcggtgga 180
actttccagg gcctcgagtc taagctcgac tacatcaagg gacttggatt cgatgccatt 240
tggattaccc ccgtcgtctc aaacaaggct gctggatacc atggctactg ggccgaggac 300
ttgtatgccg tcaactcaaa ctacggcact gctgccgact tgaagagctt ggttgccgct 360
gcccatgcca agggcatcta catgatggtc gacgttgtcg caaaccacat gggtcctgga 420
gcaatcacaa acaaccgccc tgaacctctc aaccaggctt catcatacca ccctccttgc 480

```
aacatcgact acaacaacca aaccagtgtc gaggtatgtc aaatagccgg actccccgac     540
atctacacca ccaagagcga gatccgcacg ctcctcaaca cctgggtcaa ctggctcgta     600
aacgagtaca gcttcgacgg tgtccgcatc gacaccgtca agcacgtcga aaaggacttt     660
tggcctggct tctctgccgc taccggtgtc tacaacattg gcgaggtgtt tgacggagac     720
ccagcctacc ttgccccgta cgccaagctt atgcccggcc tcctcaacta cgcagtctac     780
tacccgatga acaactttta ccagcaaacg ggctcttccc aggcgcttgt agacatgatg     840
aacactgtca gcaacacctt ccctgaccca tctgccttgg gaaccttcct cgacaaccac     900
gacaacaagc gctggttgaa cgtcaagaac gaccagactc tgctcaagaa cgctcttgct     960
tatgtcatcc tcgcacgtgg tatccccatc ttgtactatg gtaccgagca gggatacgct    1020
ggtggtgacg acccagctaa ccgagaggat ctgtggcgca gtggcttcaa caccaatgcc    1080
aacctctacc aagccatcaa gaaactgacc gccgcccgac aggctgccgg tggtctcgca    1140
ggaaacgacc acgtccacct gtacgtcgcc gacacggctt acgcctggag ccgtgccaac    1200
ggcaacctga ttgtcctcac caccaacgct ggcggcaact ccaacaccca gcactgcttc    1260
aacacgcaaa aggcaaacgg ccgctggacc aacgtctacg gcaacggcgc caccgtctct    1320
gccgatagca acggccaaat ctgcgtctcc gtcacaaacg gcgagcccgt tgtcctcctc    1380
gccggctccg ctacccccac cactggcact accctctcca cccgcaccgc cactgccacc    1440
gccacaccaa ccgcatgccc caccgccgtc tccgtctcct tcacccaccg cgtcaccact    1500
gttcccggtg acaccatcaa aatcactggc aacacggccc agctaggtaa ctggactccc    1560
gccaacggcc ttgccttgtc cgcagctagc tacacatcca gcaaccctat ctggaccatt    1620
accgtgcccc tggccgctgg atcctccatc tcgtacaagt ttgtcaagat tgacagtgga    1680
ggaactgtca cctgggagag tgaccccaac aggtcataca ctgcgccgag ctgccaggcg    1740
agtgccagtg tgaacagctc atggcaatag                                     1770

<210> 540
<211> 589
<212> PRT
<213> Cochliobolus heterostrophus ATCC 48331

<220>
<221> SIGNAL
<222> (1)...(20)

<221> DOMAIN
<222> (33)...(373)
<223> Alpha amylase, catalytic domain

<221> DOMAIN
<222> (489)...(586)
<223> Starch binding domain

<400> 540
Met Leu Leu Leu Asn Ile Phe Thr Thr Leu Phe Phe Tyr Ile Thr Cys
 1               5                  10                  15
Ile Val Ser Ala Ala Asp Thr Asn Ala Trp Lys Ser Arg Ser Ile Tyr
            20                  25                  30
Phe Val Leu Thr Asp Arg Ile Ala Arg Asn Ser Ser Asp Thr Gly Gly
        35                  40                  45
Ser Ala Cys Ser Asp Leu Gly Asn Tyr Cys Gly Gly Thr Phe Gln Gly
    50                  55                  60
Leu Glu Ser Lys Leu Asp Tyr Ile Lys Gly Leu Gly Phe Asp Ala Ile
65                  70                  75                  80
Trp Ile Thr Pro Val Val Ser Asn Lys Ala Ala Gly Tyr His Gly Tyr
                85                  90                  95
Trp Ala Glu Asp Leu Tyr Ala Val Asn Ser Asn Tyr Gly Thr Ala Ala
            100                 105                 110
Asp Leu Lys Ser Leu Val Ala Ala Ala His Ala Lys Gly Ile Tyr Met
        115                 120                 125
Met Val Asp Val Val Ala Asn His Met Gly Pro Gly Ala Ile Thr Asn
    130                 135                 140
Asn Arg Pro Glu Pro Leu Asn Gln Ala Ser Ser Tyr His Pro Pro Cys
145                 150                 155                 160
Asn Ile Asp Tyr Asn Asn Gln Thr Ser Val Glu Val Cys Gln Ile Ala
                165                 170                 175
Gly Leu Pro Asp Ile Tyr Thr Thr Lys Ser Glu Ile Arg Thr Leu Leu
            180                 185                 190
Asn Thr Trp Val Asn Trp Leu Val Asn Glu Tyr Ser Phe Asp Gly Val
        195                 200                 205
Arg Ile Asp Thr Val Lys His Val Glu Lys Asp Phe Trp Pro Gly Phe
```

```
    210                 215                 220
Ser Ala Ala Thr Gly Val Tyr Asn Ile Gly Glu Val Phe Asp Gly Asp
225                 230                 235                 240
Pro Ala Tyr Leu Ala Pro Tyr Ala Lys Leu Met Pro Gly Leu Leu Asn
                245                 250                 255
Tyr Ala Val Tyr Tyr Pro Met Asn Asn Phe Tyr Gln Gln Thr Gly Ser
            260                 265                 270
Ser Gln Ala Leu Val Asp Met Met Asn Thr Val Ser Asn Thr Phe Pro
        275                 280                 285
Asp Pro Ser Ala Leu Gly Thr Phe Leu Asp Asn His Asp Asn Lys Arg
    290                 295                 300
Trp Leu Asn Val Lys Asn Asp Gln Thr Leu Leu Lys Asn Ala Leu Ala
305                 310                 315                 320
Tyr Val Ile Leu Ala Arg Gly Ile Pro Ile Leu Tyr Tyr Gly Thr Glu
                325                 330                 335
Gln Gly Tyr Ala Gly Gly Asp Asp Pro Ala Asn Arg Glu Asp Leu Trp
            340                 345                 350
Arg Ser Gly Phe Asn Thr Asn Ala Asn Leu Tyr Gln Ala Ile Lys Lys
        355                 360                 365
Leu Thr Ala Ala Arg Gln Ala Ala Gly Gly Leu Ala Gly Asn Asp His
    370                 375                 380
Val His Leu Tyr Val Ala Asp Thr Ala Tyr Ala Trp Ser Arg Ala Asn
385                 390                 395                 400
Gly Asn Leu Ile Val Leu Thr Thr Asn Ala Gly Gly Asn Ser Asn Thr
                405                 410                 415
Gln His Cys Phe Asn Thr Gln Lys Ala Asn Gly Arg Trp Thr Asn Val
            420                 425                 430
Tyr Gly Asn Gly Ala Thr Val Ser Ala Asp Ser Asn Gly Gln Ile Cys
        435                 440                 445
Val Ser Val Thr Asn Gly Glu Pro Val Val Leu Leu Ala Gly Ser Ala
    450                 455                 460
Thr Pro Thr Thr Gly Thr Thr Leu Ser Thr Arg Thr Ala Thr Ala Thr
465                 470                 475                 480
Ala Thr Pro Thr Ala Cys Pro Thr Ala Val Ser Val Ser Phe Thr His
                485                 490                 495
Arg Val Thr Thr Val Pro Gly Asp Thr Ile Lys Ile Thr Gly Asn Thr
            500                 505                 510
Ala Gln Leu Gly Asn Trp Thr Pro Ala Asn Gly Leu Ala Leu Ser Ala
        515                 520                 525
Ala Ser Tyr Thr Ser Ser Asn Pro Ile Trp Thr Ile Thr Val Pro Leu
    530                 535                 540
Ala Ala Gly Ser Ser Ile Ser Tyr Lys Phe Val Lys Ile Asp Ser Gly
545                 550                 555                 560
Gly Thr Val Thr Trp Glu Ser Asp Pro Asn Arg Ser Tyr Thr Ala Pro
                565                 570                 575
Ser Cys Gln Ala Ser Ala Ser Val Asn Ser Ser Trp Gln
            580                 585
```

<210> 541
<211> 2037
<212> DNA
<213> Fungi

<220>
<221> CDS
<222> (1)...(109)
<223> Exon

<221> CDS
<222> (155)...(2037)
<223> Exon

<400> 541

```
atg aat gtt acg ctg cat tgt ttg agc gga gag tcc gtt atg tgc ggc       48
Met Asn Val Thr Leu His Cys Leu Ser Gly Glu Ser Val Met Cys Gly
 1               5                  10                  15

gtt gat cag aga aag gat ctg atc tcc aca agt aac ggg gtg atg tcg       96
```

```
Val Asp Gln Arg Lys Asp Leu Ile Ser Thr Ser Asn Gly Val Met Ser
            20                  25                  30

gcc gat aag atg g gtaaataata aacacctcgt cttcccctct caacaatatt        149
Ala Asp Lys Met
         35

ccaag cc cat cac tcc cat ttt acc atc atc tct tct att ctc tcc ctc     198
     Ala His His Ser His Phe Thr Ile Ile Ser Ser Ile Leu Ser Leu
                 40                  45                  50

tcc gtt cat cca gcc atc atg ggc tcc atc atc tgg gac gat gaa cac      246
Ser Val His Pro Ala Ile Met Gly Ser Ile Ile Trp Asp Asp Glu His
            55                  60                  65

caa tac ata ggc aaa cag ccc tgg tgg aaa gcc gca tcc ttc tat caa      294
Gln Tyr Ile Gly Lys Gln Pro Trp Trp Lys Ala Ala Ser Phe Tyr Gln
         70                  75                  80

gtc tac cca gcc agc ttc aaa gac tcc aac aac gac ggc tgg ggc gat      342
Val Tyr Pro Ala Ser Phe Lys Asp Ser Asn Asn Asp Gly Trp Gly Asp
     85                  90                  95

tta ccc ggt ctc atc tcc aag ctc gat tat cta agc gat ctt ggc atc      390
Leu Pro Gly Leu Ile Ser Lys Leu Asp Tyr Leu Ser Asp Leu Gly Ile
100                 105                 110                 115

gac gtc gta tgg gta tct ccc att ttc gag agt cct cag aaa gac atg      438
Asp Val Val Trp Val Ser Pro Ile Phe Glu Ser Pro Gln Lys Asp Met
                120                 125                 130

ggc tac gac gtc tca gac tat caa aaa atc tac gaa ccc tac ggc agt      486
Gly Tyr Asp Val Ser Asp Tyr Gln Lys Ile Tyr Glu Pro Tyr Gly Ser
            135                 140                 145

gtt gaa gat gtc gat att ctc gtc aaa gaa tgc cat gcg cga ggt ctc      534
Val Glu Asp Val Asp Ile Leu Val Lys Glu Cys His Ala Arg Gly Leu
        150                 155                 160

aag att att ctc gat ctt gtt gtg aat cac acg tcg ata gag cac gag      582
Lys Ile Ile Leu Asp Leu Val Val Asn His Thr Ser Ile Glu His Glu
    165                 170                 175

tgg ttc aaa gaa tcg agg tct tca aag gac aat ccc aag agg gat tgg      630
Trp Phe Lys Glu Ser Arg Ser Ser Lys Asp Asn Pro Lys Arg Asp Trp
180                 185                 190                 195

tat atc tgg aaa cca gcg cgg tat gat gag aac ggc gtg aga tat cct      678
Tyr Ile Trp Lys Pro Ala Arg Tyr Asp Glu Asn Gly Val Arg Tyr Pro
                200                 205                 210

cct acg aac tgg agg gga tat ttt gcg tgc tcg acg tgg act tgg gac      726
Pro Thr Asn Trp Arg Gly Tyr Phe Ala Cys Ser Thr Trp Thr Trp Asp
            215                 220                 225

gag cac acg caa gaa tac tat ctt cat ctt tat gcg cct gat caa cct      774
Glu His Thr Gln Glu Tyr Tyr Leu His Leu Tyr Ala Pro Asp Gln Pro
        230                 235                 240

gat ttg aat tgg gac aat gat gag tgt cga gaa gcg att tac gaa aat      822
Asp Leu Asn Trp Asp Asn Asp Glu Cys Arg Glu Ala Ile Tyr Glu Asn
    245                 250                 255

act atg cgc ttt tgg ctt gat cgc ggc gtt gat ggg ttt cgc att gat      870
Thr Met Arg Phe Trp Leu Asp Arg Gly Val Asp Gly Phe Arg Ile Asp
260                 265                 270                 275

acc gtt aac aaa tac tcc aag agg cga gac ttt gtc gac gcg cct atc      918
```

```
Thr Val Asn Lys Tyr Ser Lys Arg Arg Asp Phe Val Asp Ala Pro Ile
                280                 285                 290

acg gat ccc tcg tcg cct cat cag cct gcg cct gaa atg tgg tgc aac      966
Thr Asp Pro Ser Ser Pro His Gln Pro Ala Pro Glu Met Trp Cys Asn
            295                 300                 305

gga ccg cgc att cat gag ttt atc cat gaa atg tat gag aaa gtc ctt     1014
Gly Pro Arg Ile His Glu Phe Ile His Glu Met Tyr Glu Lys Val Leu
        310                 315                 320

ggg cct tat aat gcg gtt tca gtg ggt gag ctt tca aac acg cct aat     1062
Gly Pro Tyr Asn Ala Val Ser Val Gly Glu Leu Ser Asn Thr Pro Asn
    325                 330                 335

cct tca caa gtc att ccc tat gtc tca gcc gct gcg aaa gag ttg gac     1110
Pro Ser Gln Val Ile Pro Tyr Val Ser Ala Ala Ala Lys Glu Leu Asp
340                 345                 350                 355

atg gtg ttt gag ttc tcc atg atc aga ctt ggt aat ggc aat ggg ttt     1158
Met Val Phe Glu Phe Ser Met Ile Arg Leu Gly Asn Gly Asn Gly Phe
                360                 365                 370

ggt gac aag tat att tac gaa ccg ttc ccg tta tcg aag ctc aag gcg     1206
Gly Asp Lys Tyr Ile Tyr Glu Pro Phe Pro Leu Ser Lys Leu Lys Ala
            375                 380                 385

ttt gtg tat aag tgg cag agc ttt att gag ggt aca gac gcg tgg gca     1254
Phe Val Tyr Lys Trp Gln Ser Phe Ile Glu Gly Thr Asp Ala Trp Ala
        390                 395                 400

aca gct ttc tgc gag aac cac gac aat ggc cgt gcg gtc gac cgt ttc     1302
Thr Ala Phe Cys Glu Asn His Asp Asn Gly Arg Ala Val Asp Arg Phe
    405                 410                 415

ggt gac gcg tca acg cca gag tgg tgg cta aag tca gca aag act ttg     1350
Gly Asp Ala Ser Thr Pro Glu Trp Trp Leu Lys Ser Ala Lys Thr Leu
420                 425                 430                 435

gct acc tgg caa acg acg ctc acg ggc act tta ttc cta tac caa gga     1398
Ala Thr Trp Gln Thr Thr Leu Thr Gly Thr Leu Phe Leu Tyr Gln Gly
                440                 445                 450

caa gaa att ggc atg acc aac atg ccc aag tcg tgg gga att gaa gaa     1446
Gln Glu Ile Gly Met Thr Asn Met Pro Lys Ser Trp Gly Ile Glu Glu
            455                 460                 465

tac aag gat atc gaa agc agc aac ttc tac gct gaa gcc gtc gcg tca     1494
Tyr Lys Asp Ile Glu Ser Ser Asn Phe Tyr Ala Glu Ala Val Ala Ser
        470                 475                 480

gga gac gag aag cgc gtt aga gac acc atg cac ggt ctt caa atc atg     1542
Gly Asp Glu Lys Arg Val Arg Asp Thr Met His Gly Leu Gln Ile Met
    485                 490                 495

gcg cgt gat cac tcg cgt att ccc ttc caa tgg gat gat tca ccc aac     1590
Ala Arg Asp His Ser Arg Ile Pro Phe Gln Trp Asp Asp Ser Pro Asn
500                 505                 510                 515

gct ggc ttt gcc gac gcg tct gcg aag cca tgg atg aga gtc cac gat     1638
Ala Gly Phe Ala Asp Ala Ser Ala Lys Pro Trp Met Arg Val His Asp
                520                 525                 530

gac tat cgc gat atc aac gtc gct aag cag gtc aaa gac cca aac tcc     1686
Asp Tyr Arg Asp Ile Asn Val Ala Lys Gln Val Lys Asp Pro Asn Ser
            535                 540                 545

att ctc agt ttc tac aaa gcc atg ctg cgc ctt cga aaa gaa tat caa     1734
```

```
Ile Leu Ser Phe Tyr Lys Ala Met Leu Arg Leu Arg Lys Glu Tyr Gln
        550                 555                 560

gat atc ttt gta ttt ggt tct ttc aaa ctt cta gat cct gaa gat gaa     1782
Asp Ile Phe Val Phe Gly Ser Phe Lys Leu Leu Asp Pro Glu Asp Glu
    565                 570                 575

tcg ctg ttt tgt tat gtg aaa gag agt gtt gca cct gtt aag gga cga     1830
Ser Leu Phe Cys Tyr Val Lys Glu Ser Val Ala Pro Val Lys Gly Arg
580                 585                 590                 595

cgg ggt gag aag agg aag gct gtt gtt gtg ttg aat atg agt cga gag     1878
Arg Gly Glu Lys Arg Lys Ala Val Val Val Leu Asn Met Ser Arg Glu
                600                 605                 610

gat aga ctt ggc ccg gat gtt cct act gtc ttg gga tgc tcg cat gag     1926
Asp Arg Leu Gly Pro Asp Val Pro Thr Val Leu Gly Cys Ser His Glu
            615                 620                 625

gat gtg agg ttg atg gcg tat acg agg gag act gcg ggc gga ttg aaa     1974
Asp Val Arg Leu Met Ala Tyr Thr Arg Glu Thr Ala Gly Gly Leu Lys
        630                 635                 640

gtg gga cgg ccg atg ctt gcg ggc tgg gag agt cgg gtt tat gtt aac     2022
Val Gly Arg Pro Met Leu Ala Gly Trp Glu Ser Arg Val Tyr Val Asn
    645                 650                 655

ttt gag ctg gaa tag                                                2037
Phe Glu Leu Glu  *
660


<210> 542
<211> 36
<212> PRT
<213> Fungi

<400> 542
Met Asn Val Thr Leu His Cys Leu Ser Gly Glu Ser Val Met Cys Gly
 1               5                  10                  15
Val Asp Gln Arg Lys Asp Leu Ile Ser Thr Ser Asn Gly Val Met Ser
            20                  25                  30
Ala Asp Lys Met
        35

<210> 543
<211> 627
<212> PRT
<213> Fungi

<400> 543
Ala His His Ser His Phe Thr Ile Ile Ser Ser Ile Leu Ser Leu Ser
 1               5                  10                  15
Val His Pro Ala Ile Met Gly Ser Ile Ile Trp Asp Asp Glu His Gln
            20                  25                  30
Tyr Ile Gly Lys Gln Pro Trp Trp Lys Ala Ala Ser Phe Tyr Gln Val
        35                  40                  45
Tyr Pro Ala Ser Phe Lys Asp Ser Asn Asn Asp Gly Trp Gly Asp Leu
    50                  55                  60
Pro Gly Leu Ile Ser Lys Leu Asp Tyr Leu Ser Asp Leu Gly Ile Asp
65                  70                  75                  80
Val Val Trp Val Ser Pro Ile Phe Glu Ser Pro Gln Lys Asp Met Gly
                85                  90                  95
Tyr Asp Val Ser Asp Tyr Gln Lys Ile Tyr Glu Pro Tyr Gly Ser Val
            100                 105                 110
Glu Asp Val Asp Ile Leu Val Lys Glu Cys His Ala Arg Gly Leu Lys
        115                 120                 125
Ile Ile Leu Asp Leu Val Val Asn His Thr Ser Ile Glu His Glu Trp
```

```
    130                 135                 140
Phe Lys Glu Ser Arg Ser Ser Lys Asp Asn Pro Lys Arg Asp Trp Tyr
145                 150                 155                 160
Ile Trp Lys Pro Ala Arg Tyr Asp Glu Asn Gly Val Arg Tyr Pro Pro
                165                 170                 175
Thr Asn Trp Arg Gly Tyr Phe Ala Cys Ser Thr Trp Thr Trp Asp Glu
            180                 185                 190
His Thr Gln Glu Tyr Tyr Leu His Leu Tyr Ala Pro Asp Gln Pro Asp
        195                 200                 205
Leu Asn Trp Asp Asn Asp Glu Cys Arg Glu Ala Ile Tyr Glu Asn Thr
    210                 215                 220
Met Arg Phe Trp Leu Asp Arg Gly Val Asp Gly Phe Arg Ile Asp Thr
225                 230                 235                 240
Val Asn Lys Tyr Ser Lys Arg Arg Asp Phe Val Asp Ala Pro Ile Thr
                245                 250                 255
Asp Pro Ser Ser Pro His Gln Pro Ala Pro Glu Met Trp Cys Asn Gly
            260                 265                 270
Pro Arg Ile His Glu Phe Ile His Glu Met Tyr Glu Lys Val Leu Gly
        275                 280                 285
Pro Tyr Asn Ala Val Ser Val Gly Glu Leu Ser Asn Thr Pro Asn Pro
    290                 295                 300
Ser Gln Val Ile Pro Tyr Val Ser Ala Ala Ala Lys Glu Leu Asp Met
305                 310                 315                 320
Val Phe Glu Phe Ser Met Ile Arg Leu Gly Asn Gly Asn Gly Phe Gly
                325                 330                 335
Asp Lys Tyr Ile Tyr Glu Pro Phe Pro Leu Ser Lys Leu Lys Ala Phe
            340                 345                 350
Val Tyr Lys Trp Gln Ser Phe Ile Glu Gly Thr Asp Ala Trp Ala Thr
        355                 360                 365
Ala Phe Cys Glu Asn His Asp Asn Gly Arg Ala Val Asp Arg Phe Gly
    370                 375                 380
Asp Ala Ser Thr Pro Glu Trp Trp Leu Lys Ser Ala Lys Thr Leu Ala
385                 390                 395                 400
Thr Trp Gln Thr Thr Leu Thr Gly Thr Leu Phe Leu Tyr Gln Gly Gln
                405                 410                 415
Glu Ile Gly Met Thr Asn Met Pro Lys Ser Trp Gly Ile Glu Glu Tyr
            420                 425                 430
Lys Asp Ile Glu Ser Ser Asn Phe Tyr Ala Glu Ala Val Ala Ser Gly
        435                 440                 445
Asp Glu Lys Arg Val Arg Asp Thr Met His Gly Leu Gln Ile Met Ala
    450                 455                 460
Arg Asp His Ser Arg Ile Pro Phe Gln Trp Asp Asp Ser Pro Asn Ala
465                 470                 475                 480
Gly Phe Ala Asp Ala Ser Ala Lys Pro Trp Met Arg Val His Asp Asp
                485                 490                 495
Tyr Arg Asp Ile Asn Val Ala Lys Gln Val Lys Asp Pro Asn Ser Ile
            500                 505                 510
Leu Ser Phe Tyr Lys Ala Met Leu Arg Leu Arg Lys Glu Tyr Gln Asp
        515                 520                 525
Ile Phe Val Phe Gly Ser Phe Lys Leu Leu Asp Pro Glu Asp Glu Ser
    530                 535                 540
Leu Phe Cys Tyr Val Lys Glu Ser Val Ala Pro Val Lys Gly Arg Arg
545                 550                 555                 560
Gly Glu Lys Arg Lys Ala Val Val Val Leu Asn Met Ser Arg Glu Asp
                565                 570                 575
Arg Leu Gly Pro Asp Val Pro Thr Val Leu Gly Cys Ser His Glu Asp
            580                 585                 590
Val Arg Leu Met Ala Tyr Thr Arg Glu Thr Ala Gly Gly Leu Lys Val
        595                 600                 605
Gly Arg Pro Met Leu Ala Gly Trp Glu Ser Arg Val Tyr Val Asn Phe
    610                 615                 620
Glu Leu Glu
625

<210> 544
<211> 1992
<212> DNA
<213> Fungi
```

<400> 544
atgaatgtta cgctgcattg tttgagcgga gagtccgtta tgtgcggcgt tgatcagaga 60
aaggatctga tctccacaag taacggggtg atgtcggccg ataagatggc ccatcactcc 120
cattttacca tcatctcttc tattctctcc ctctccgttc atccagccat catgggctcc 180
atcatctggg acgatgaaca ccaatacata ggcaaacagc cctggtggaa agccgcatcc 240
ttctatcaag tctacccagc cagcttcaaa gactccaaca acgacggctg gggcgattta 300
cccggtctca tctccaagct cgattatcta agcgatcttg gcatcgacgt cgtatgggta 360
tctcccattt tcgagagtcc tcagaaagac atgggctacg acgtctcaga ctatcaaaaa 420
atctacgaac cctacggcag tgttgaagat gtcgatattc tcgtcaaaga atgccatgcg 480
cgaggtctca agattattct cgatcttgtt gtgaatcaca cgtcgataga gcacgagtgg 540
ttcaaagaat cgaggtcttc aaaggacaat cccaagaggg attggtatat ctggaaacca 600
gcgcggtatg atgagaacgg cgtgagatat cctcctacga actggagggg atattttgcg 660
tgctcgacgt ggacttggga cgagcacacg caagaatact atcttcatct ttatgcgcct 720
gatcaacctg atttgaattg ggacaatgat gagtgtcgag aagcgattta cgaaaatact 780
atgcgctttt ggcttgatcg cggcgttgat gggtttcgca ttgataccgt taacaaatac 840
tccaagaggc gagactttgt cgacgcgcct atcacggatc cctcgtcgcc tcatcagcct 900
gcgcctgaaa tgtggtgcaa cggaccgcgc attcatgagt ttatccatga aatgtatgag 960
aaagtccttg ggccttataa tgcggtttca gtgggtgagc tttcaaacac gcctaatcct 1020
tcacaagtca ttccctatgt ctcagccgct gcgaaagagt tggacatggt gtttgagttc 1080
tccatgatca gacttggtaa tggcaatggg tttggtgaca agtatattta cgaaccgttc 1140
ccgttatcga agctcaaggc gtttgtgtat aagtggcaga gctttattga gggtacagac 1200
gcgtgggcaa cagctttctg cgagaaccac gacaatggcc gtgcggtcga ccgtttcggt 1260
gacgcgtcaa cgccagagtg gtggctaaag tcagcaaaga ctttggctac ctggcaaacg 1320
acgctcacgg gcactttatt cctataccaa ggacaagaaa ttggcatgac caacatgccc 1380
aagtcgtggg gaattgaaga atacaaggat atcgaaagca gcaacttcta cgctgaagcc 1440
gtcgcgtcag gagacgagaa gcgcgttaga gacaccatgc acggtcttca aatcatggcg 1500
cgtgatcact cgcgtattcc cttccaatgg gatgattcac ccaacgctgg ctttgccgac 1560
gcgtctgcga agccatggat gagagtccac gatgactatc gcgatatcaa cgtcgctaag 1620
caggtcaaag acccaaactc cattctcagt ttctacaaag ccatgctgcg ccttcgaaaa 1680
gaatatcaag atatctttgt atttggttct ttcaaacttc tagatcctga agatgaatcg 1740
ctgttttgtt atgtgaaaga gagtgttgca cctgttaagg gacgacgggg tgagaagagg 1800
aaggctgttg ttgtgttgaa tatgagtcga gaggatagac ttggcccgga tgttcctact 1860
gtcttgggat gctcgcatga ggatgtgagg ttgatggcgt atacgaggga gactgcgggc 1920
ggattgaaag tgggacggcc gatgcttgcg ggctgggaga gtcgggttta tgttaacttt 1980
gagctggaat ag 1992

<210> 545
<211> 663
<212> PRT
<213> Fungi

<220>
<221> DOMAIN
<222> (83)...(505)
<223> Alpha amylase, catalytic domain

<400> 545
Met Asn Val Thr Leu His Cys Leu Ser Gly Glu Ser Val Met Cys Gly
 1   5   10   15
Val Asp Gln Arg Lys Asp Leu Ile Ser Thr Ser Asn Gly Val Met Ser
            20   25   30
Ala Asp Lys Met Ala His His Ser His Phe Thr Ile Ile Ser Ser Ile
        35   40   45
Leu Ser Leu Ser Val His Pro Ala Ile Met Gly Ser Ile Ile Trp Asp
    50   55   60
Asp Glu His Gln Tyr Ile Gly Lys Gln Pro Trp Trp Lys Ala Ala Ser
65   70   75   80
Phe Tyr Gln Val Tyr Pro Ala Ser Phe Lys Asp Ser Asn Asn Asp Gly
                85   90   95
Trp Gly Asp Leu Pro Gly Leu Ile Ser Lys Leu Asp Tyr Leu Ser Asp
            100   105   110
Leu Gly Ile Asp Val Val Trp Val Ser Pro Ile Phe Glu Ser Pro Gln
        115   120   125
Lys Asp Met Gly Tyr Asp Val Ser Asp Tyr Gln Lys Ile Tyr Glu Pro
    130   135   140
Tyr Gly Ser Val Glu Asp Val Asp Ile Leu Val Lys Glu Cys His Ala
145   150   155   160

```
Arg Gly Leu Lys Ile Ile Leu Asp Leu Val Val Asn His Thr Ser Ile
                165                 170                 175
Glu His Glu Trp Phe Lys Glu Ser Arg Ser Ser Lys Asp Asn Pro Lys
            180                 185                 190
Arg Asp Trp Tyr Ile Trp Lys Pro Ala Arg Tyr Asp Glu Asn Gly Val
        195                 200                 205
Arg Tyr Pro Pro Thr Asn Trp Arg Gly Tyr Phe Ala Cys Ser Thr Trp
    210                 215                 220
Thr Trp Asp Glu His Thr Gln Glu Tyr Tyr Leu His Leu Tyr Ala Pro
225                 230                 235                 240
Asp Gln Pro Asp Leu Asn Trp Asp Asn Asp Glu Cys Arg Glu Ala Ile
                245                 250                 255
Tyr Glu Asn Thr Met Arg Phe Trp Leu Asp Arg Gly Val Asp Gly Phe
            260                 265                 270
Arg Ile Asp Thr Val Asn Lys Tyr Ser Lys Arg Arg Asp Phe Val Asp
        275                 280                 285
Ala Pro Ile Thr Asp Pro Ser Ser Pro His Gln Pro Ala Pro Glu Met
    290                 295                 300
Trp Cys Asn Gly Pro Arg Ile His Glu Phe Ile His Glu Met Tyr Glu
305                 310                 315                 320
Lys Val Leu Gly Pro Tyr Asn Ala Val Ser Val Gly Glu Leu Ser Asn
                325                 330                 335
Thr Pro Asn Pro Ser Gln Val Ile Pro Tyr Val Ser Ala Ala Ala Lys
            340                 345                 350
Glu Leu Asp Met Val Phe Glu Phe Ser Met Ile Arg Leu Gly Asn Gly
        355                 360                 365
Asn Gly Phe Gly Asp Lys Tyr Ile Tyr Glu Pro Phe Pro Leu Ser Lys
    370                 375                 380
Leu Lys Ala Phe Val Tyr Lys Trp Gln Ser Phe Ile Glu Gly Thr Asp
385                 390                 395                 400
Ala Trp Ala Thr Ala Phe Cys Glu Asn His Asp Asn Gly Arg Ala Val
                405                 410                 415
Asp Arg Phe Gly Asp Ala Ser Thr Pro Glu Trp Trp Leu Lys Ser Ala
            420                 425                 430
Lys Thr Leu Ala Thr Trp Gln Thr Thr Leu Thr Gly Thr Leu Phe Leu
        435                 440                 445
Tyr Gln Gly Gln Glu Ile Gly Met Thr Asn Met Pro Lys Ser Trp Gly
    450                 455                 460
Ile Glu Glu Tyr Lys Asp Ile Glu Ser Ser Asn Phe Tyr Ala Glu Ala
465                 470                 475                 480
Val Ala Ser Gly Asp Glu Lys Arg Val Arg Asp Thr Met His Gly Leu
                485                 490                 495
Gln Ile Met Ala Arg Asp His Ser Arg Ile Pro Phe Gln Trp Asp Asp
            500                 505                 510
Ser Pro Asn Ala Gly Phe Ala Asp Ala Ser Ala Lys Pro Trp Met Arg
        515                 520                 525
Val His Asp Asp Tyr Arg Asp Ile Asn Val Ala Lys Gln Val Lys Asp
    530                 535                 540
Pro Asn Ser Ile Leu Ser Phe Tyr Lys Ala Met Leu Arg Leu Arg Lys
545                 550                 555                 560
Glu Tyr Gln Asp Ile Phe Val Phe Gly Ser Phe Lys Leu Leu Asp Pro
                565                 570                 575
Glu Asp Glu Ser Leu Phe Cys Tyr Val Lys Glu Ser Val Ala Pro Val
            580                 585                 590
Lys Gly Arg Arg Gly Glu Lys Arg Lys Ala Val Val Val Leu Asn Met
        595                 600                 605
Ser Arg Glu Asp Arg Leu Gly Pro Asp Val Pro Thr Val Leu Gly Cys
    610                 615                 620
Ser His Glu Asp Val Arg Leu Met Ala Tyr Thr Arg Glu Thr Ala Gly
625                 630                 635                 640
Gly Leu Lys Val Gly Arg Pro Met Leu Ala Gly Trp Glu Ser Arg Val
                645                 650                 655
Tyr Val Asn Phe Glu Leu Glu
            660

<210> 546
<211> 2261
<212> DNA
```

<213> Fungi

<220>
<221> CDS
<222> (1)...(328)
<223> Exon

<221> CDS
<222> (381)...(593)
<223> Exon

<221> CDS
<222> (648)...(803)
<223> Exon

<221> CDS
<222> (887)...(1025)
<223> Exon

<221> CDS
<222> (1079)...(2261)
<223> Exon

<400> 546

```
atg gct tct tct ctg ctt tcc tcc ctc agc tcc atc agc act ttc aac       48
Met Ala Ser Ser Leu Leu Ser Ser Leu Ser Ser Ile Ser Thr Phe Asn
 1               5                  10                  15

agc act cag atc ctc caa gct gtt caa cct att atg cat aca ttc gct       96
Ser Thr Gln Ile Leu Gln Ala Val Gln Pro Ile Met His Thr Phe Ala
            20                  25                  30

ccc ctc ctt ttg gtt ggc agt ctt gcc ctc caa act atc tat gct ttg      144
Pro Leu Leu Leu Val Gly Ser Leu Ala Leu Gln Thr Ile Tyr Ala Leu
         35                  40                  45

cct ggc ccc agt aaa gtt gag caa cgc gcc gaa atc cta aaa cgc tca      192
Pro Gly Pro Ser Lys Val Glu Gln Arg Ala Glu Ile Leu Lys Arg Ser
     50                  55                  60

gtg gat tct ttt att gcc acc gaa agt cct att gct ttc aga aat ctt      240
Val Asp Ser Phe Ile Ala Thr Glu Ser Pro Ile Ala Phe Arg Asn Leu
 65                  70                  75                  80

tta tgc aat atc ggc gca gat gga gct tgt gct tca ggg gca gct tcg      288
Leu Cys Asn Ile Gly Ala Asp Gly Ala Cys Ala Ser Gly Ala Ala Ser
                85                  90                  95

ggt att gtg gtt gct tcg cct gac aaa gtg agc cca gat t gttagtattt     338
Gly Ile Val Val Ala Ser Pro Asp Lys Val Ser Pro Asp
            100                 105

catgatgccg tttgaataca tggaattctg atatgtaaac ag at ttt tat act tgg    394
                                              Tyr Phe Tyr Thr Trp
                                              110

act cga gac tcc gca ttg gta ttc aaa gcc ctt gtt gat acg ttt atc      442
Thr Arg Asp Ser Ala Leu Val Phe Lys Ala Leu Val Asp Thr Phe Ile
115                 120                 125                 130

agc agc agc tat tct gca tca ctt caa cag gaa att gaa aat tac atc      490
Ser Ser Ser Tyr Ser Ala Ser Leu Gln Gln Glu Ile Glu Asn Tyr Ile
                135                 140                 145

agc tcc caa gct gga ctt caa acc gtg tca aac cct tca ggt ggt tta      538
Ser Ser Gln Ala Gly Leu Gln Thr Val Ser Asn Pro Ser Gly Gly Leu
            150                 155                 160
```

```
tca tcc ggt ggt ctg ggc gag ccc aaa ttt aat gcg gat aag tcg gca      586
Ser Ser Gly Gly Leu Gly Glu Pro Lys Phe Asn Ala Asp Lys Ser Ala
        165                 170                 175

ttc aca g gtacgaaaat tcgaaaactc tatcctgatg gtgactgaca tgtgaactat     643
Phe Thr
    180

acag gg tct tgg gga aga cct caa cga gat gga ccc gct tta cgt gct      691
    Gly Ser Trp Gly Arg Pro Gln Arg Asp Gly Pro Ala Leu Arg Ala
                    185                 190                 195

act gct atg att acc tat tca aaa tgg ttg att gcc aac gga tac acc      739
Thr Ala Met Ile Thr Tyr Ser Lys Trp Leu Ile Ala Asn Gly Tyr Thr
                200                 205                 210

tcc act gtg caa act att gtt tgg ccg atc atc aga aat gat ctg tct      787
Ser Thr Val Gln Thr Ile Val Trp Pro Ile Ile Arg Asn Asp Leu Ser
            215                 220                 225

tat gtt act caa tat t gtaagtcaac gggaagaaaa ccaaattagt ctaaaaatag    843
Tyr Val Thr Gln Tyr
        230

atgaccctct tttcttggag cttttgatta acttaattca tag gg aac caa act       897
                                                Trp Asn Gln Thr
                                                    235

ggc ttt gat ctc tgg gaa gag gtt caa ggc tca agt ttc ttc aca gtc      945
Gly Phe Asp Leu Trp Glu Glu Val Gln Gly Ser Ser Phe Phe Thr Val
            240                 245                 250

gct gct caa cat cgt gct ctt gtt gaa ggt agt gca ctg gcc gct caa      993
Ala Ala Gln His Arg Ala Leu Val Glu Gly Ser Ala Leu Ala Ala Gln
        255                 260                 265

att ggt caa agc tgt acc tat tgt gat tct ca gtaagcacat atgaaagaat    1045
Ile Gly Gln Ser Cys Thr Tyr Cys Asp Ser Gln
    270                 275

atattcacat gcaggtacta atacattttg cag a gcc cca cag gtt ctt tgt      1097
                                    Ala Pro Gln Val Leu Cys
                                    280                 285

ttc ctc caa agc ttc tgg agt tca tcc tcg gga tac atc att tcc aac     1145
Phe Leu Gln Ser Phe Trp Ser Ser Ser Ser Gly Tyr Ile Ile Ser Asn
                290                 295                 300

atc aac caa aac agc gga aga aac gga aaa gat gca aac agt atc ttg     1193
Ile Asn Gln Asn Ser Gly Arg Asn Gly Lys Asp Ala Asn Ser Ile Leu
            305                 310                 315

agt tct atc cag acc ttt gat cct acc gct gca tgt gat gcc acc acg     1241
Ser Ser Ile Gln Thr Phe Asp Pro Thr Ala Ala Cys Asp Ala Thr Thr
        320                 325                 330

ttc caa cca tgc tca gac cgt gct ttg gca aat cac aaa gtt gtg act     1289
Phe Gln Pro Cys Ser Asp Arg Ala Leu Ala Asn His Lys Val Val Thr
    335                 340                 345

gat tct ttc aga tct atc tac agc atc aac tct ggg att gcc gaa ggt     1337
Asp Ser Phe Arg Ser Ile Tyr Ser Ile Asn Ser Gly Ile Ala Glu Gly
350                 355                 360                 365

gtc gcc gtt tct gtt gga cgc tac cca gaa gat tct tac tat ggt ggc     1385
Val Ala Val Ser Val Gly Arg Tyr Pro Glu Asp Ser Tyr Tyr Gly Gly
                370                 375                 380
```

```
aat cct tgg tat ctc aat acc ctt gct gct gcc gaa cag ctt tat gat     1433
Asn Pro Trp Tyr Leu Asn Thr Leu Ala Ala Ala Glu Gln Leu Tyr Asp
            385                 390                 395

gct tta tac acc tgg aat aga att ggc tct atc acc gta act tcc act     1481
Ala Leu Tyr Thr Trp Asn Arg Ile Gly Ser Ile Thr Val Thr Ser Thr
        400                 405                 410

tcc cta gca ttc ttc aaa gat ttc agc tct tcc atc aca gcc gga aca     1529
Ser Leu Ala Phe Phe Lys Asp Phe Ser Ser Ser Ile Thr Ala Gly Thr
    415                 420                 425

tac gca tcc tcc aca tca acc tac acc act ctc tac aac gcc atc aaa     1577
Tyr Ala Ser Ser Thr Ser Thr Tyr Thr Thr Leu Tyr Asn Ala Ile Lys
430                 435                 440                 445

acc tat gcc gac ggc tac gtc aat gtc gtc gca acc tac gca caa gcc     1625
Thr Tyr Ala Asp Gly Tyr Val Asn Val Val Ala Thr Tyr Ala Gln Ala
                450                 455                 460

aac ggt tcc ctc tca gaa caa ttc aac aaa gcc ggc ggc gct ccc ctc     1673
Asn Gly Ser Leu Ser Glu Gln Phe Asn Lys Ala Gly Gly Ala Pro Leu
            465                 470                 475

tca gcc tac gac ctc acc tgg tcc tat gct gcc ttc ctc acc gcc gct     1721
Ser Ala Tyr Asp Leu Thr Trp Ser Tyr Ala Ala Phe Leu Thr Ala Ala
        480                 485                 490

gcc cgt cgc gca ggt gtc gtc cca tac tca tgg ggt gag cca tcc gcc     1769
Ala Arg Arg Ala Gly Val Val Pro Tyr Ser Trp Gly Glu Pro Ser Ala
    495                 500                 505

tcc agc gtc cca gga acc tgt tct gct acc tcg gcc atc gga acc tat     1817
Ser Ser Val Pro Gly Thr Cys Ser Ala Thr Ser Ala Ile Gly Thr Tyr
510                 515                 520                 525

tct acc gca acc gcc act tcc tgg ccc gca tcg caa acc ccc agc gga     1865
Ser Thr Ala Thr Ala Thr Ser Trp Pro Ala Ser Gln Thr Pro Ser Gly
                530                 535                 540

gga gga gga agt acc ccc aca acc agc aca tcc gcc acc acc aag aca     1913
Gly Gly Gly Ser Thr Pro Thr Thr Ser Thr Ser Ala Thr Thr Lys Thr
            545                 550                 555

act tcc acc tcg acc acc act tct tca tgc gca aca gcc act tca gta     1961
Thr Ser Thr Ser Thr Thr Thr Ser Ser Cys Ala Thr Ala Thr Ser Val
        560                 565                 570

gct gtt act ttc aat gaa ttg gtc act acc tcc ttt ggc caa acg atc     2009
Ala Val Thr Phe Asn Glu Leu Val Thr Thr Ser Phe Gly Gln Thr Ile
    575                 580                 585

aaa ctc gct ggt tcc gtc tct caa ttg gga agt tgg gca ccg gct tct     2057
Lys Leu Ala Gly Ser Val Ser Gln Leu Gly Ser Trp Ala Pro Ala Ser
590                 595                 600                 605

gcg att gca ttg agc gcc gcg aag tac acc gct agt aat ccg ttg tgg     2105
Ala Ile Ala Leu Ser Ala Ala Lys Tyr Thr Ala Ser Asn Pro Leu Trp
                610                 615                 620

acc gtg act gtt aac cta cca gct ggc acg acg gta ttg tat aag ttt     2153
Thr Val Thr Val Asn Leu Pro Ala Gly Thr Thr Val Leu Tyr Lys Phe
            625                 630                 635

att aac gtg gcg agt gat ggg acg gtt acc tgg cag gcg gat cca aat     2201
Ile Asn Val Ala Ser Asp Gly Thr Val Thr Trp Gln Ala Asp Pro Asn
        640                 645                 650
```

```
aag agt tat act gtt ccg gta ggc tgt gct acg act gct act gtt agt     2249
Lys Ser Tyr Thr Val Pro Val Gly Cys Ala Thr Thr Ala Thr Val Ser
    655                 660                 665

ggt agc tgg taa                                                    2261
Gly Ser Trp  *
670


<210> 547
<211> 109
<212> PRT
<213> Fungi

<400> 547
Met Ala Ser Ser Leu Leu Ser Ser Leu Ser Ser Ile Ser Thr Phe Asn
 1               5                  10                  15
Ser Thr Gln Ile Leu Gln Ala Val Gln Pro Ile Met His Thr Phe Ala
            20                  25                  30
Pro Leu Leu Leu Val Gly Ser Leu Ala Leu Gln Thr Ile Tyr Ala Leu
        35                  40                  45
Pro Gly Pro Ser Lys Val Glu Gln Arg Ala Glu Ile Leu Lys Arg Ser
    50                  55                  60
Val Asp Ser Phe Ile Ala Thr Glu Ser Pro Ile Ala Phe Arg Asn Leu
65                  70                  75                  80
Leu Cys Asn Ile Gly Ala Asp Gly Ala Cys Ala Ser Gly Ala Ala Ser
                85                  90                  95
Gly Ile Val Val Ala Ser Pro Asp Lys Val Ser Pro Asp
            100                 105

<210> 548
<211> 71
<212> PRT
<213> Fungi

<400> 548
Tyr Phe Tyr Thr Trp Thr Arg Asp Ser Ala Leu Val Phe Lys Ala Leu
 1               5                  10                  15
Val Asp Thr Phe Ile Ser Ser Ser Tyr Ser Ala Ser Leu Gln Gln Glu
            20                  25                  30
Ile Glu Asn Tyr Ile Ser Ser Gln Ala Gly Leu Gln Thr Val Ser Asn
        35                  40                  45
Pro Ser Gly Gly Leu Ser Ser Gly Gly Leu Gly Glu Pro Lys Phe Asn
    50                  55                  60
Ala Asp Lys Ser Ala Phe Thr
65                  70

<210> 549
<211> 52
<212> PRT
<213> Fungi

<400> 549
Gly Ser Trp Gly Arg Pro Gln Arg Asp Gly Pro Ala Leu Arg Ala Thr
 1               5                  10                  15
Ala Met Ile Thr Tyr Ser Lys Trp Leu Ile Ala Asn Gly Tyr Thr Ser
            20                  25                  30
Thr Val Gln Thr Ile Val Trp Pro Ile Ile Arg Asn Asp Leu Ser Tyr
        35                  40                  45
Val Thr Gln Tyr
    50

<210> 550
<211> 47
<212> PRT
<213> Fungi

<400> 550
```

```
Trp Asn Gln Thr Gly Phe Asp Leu Trp Glu Glu Val Gln Gly Ser Ser
 1               5                  10                  15
Phe Phe Thr Val Ala Ala Gln His Arg Ala Leu Val Glu Gly Ser Ala
            20                  25                  30
Leu Ala Ala Gln Ile Gly Gln Ser Cys Thr Tyr Cys Asp Ser Gln
        35                  40                  45

<210> 551
<211> 393
<212> PRT
<213> Fungi

<400> 551
Ala Pro Gln Val Leu Cys Phe Leu Gln Ser Phe Trp Ser Ser Ser Ser
 1               5                  10                  15
Gly Tyr Ile Ile Ser Asn Ile Asn Gln Asn Ser Gly Arg Asn Gly Lys
            20                  25                  30
Asp Ala Asn Ser Ile Leu Ser Ser Ile Gln Thr Phe Asp Pro Thr Ala
        35                  40                  45
Ala Cys Asp Ala Thr Thr Phe Gln Pro Cys Ser Asp Arg Ala Leu Ala
    50                  55                  60
Asn His Lys Val Val Thr Asp Ser Phe Arg Ser Ile Tyr Ser Ile Asn
65                  70                  75                  80
Ser Gly Ile Ala Glu Gly Val Ala Val Ser Val Gly Arg Tyr Pro Glu
                85                  90                  95
Asp Ser Tyr Tyr Gly Gly Asn Pro Trp Tyr Leu Asn Thr Leu Ala Ala
            100                 105                 110
Ala Glu Gln Leu Tyr Asp Ala Leu Tyr Thr Trp Asn Arg Ile Gly Ser
        115                 120                 125
Ile Thr Val Thr Ser Thr Ser Leu Ala Phe Phe Lys Asp Phe Ser Ser
    130                 135                 140
Ser Ile Thr Ala Gly Thr Tyr Ala Ser Ser Thr Ser Thr Tyr Thr Thr
145                 150                 155                 160
Leu Tyr Asn Ala Ile Lys Thr Tyr Ala Asp Gly Tyr Val Asn Val Val
                165                 170                 175
Ala Thr Tyr Ala Gln Ala Asn Gly Ser Leu Ser Glu Gln Phe Asn Lys
            180                 185                 190
Ala Gly Gly Ala Pro Leu Ser Ala Tyr Asp Leu Thr Trp Ser Tyr Ala
        195                 200                 205
Ala Phe Leu Thr Ala Ala Ala Arg Arg Ala Gly Val Val Pro Tyr Ser
    210                 215                 220
Trp Gly Glu Pro Ser Ala Ser Ser Val Pro Gly Thr Cys Ser Ala Thr
225                 230                 235                 240
Ser Ala Ile Gly Thr Tyr Ser Thr Ala Thr Ala Thr Ser Trp Pro Ala
                245                 250                 255
Ser Gln Thr Pro Ser Gly Gly Gly Gly Ser Thr Pro Thr Thr Ser Thr
            260                 265                 270
Ser Ala Thr Thr Lys Thr Thr Ser Thr Ser Thr Thr Thr Ser Ser Cys
        275                 280                 285
Ala Thr Ala Thr Ser Val Ala Val Thr Phe Asn Glu Leu Val Thr Thr
    290                 295                 300
Ser Phe Gly Gln Thr Ile Lys Leu Ala Gly Ser Val Ser Gln Leu Gly
305                 310                 315                 320
Ser Trp Ala Pro Ala Ser Ala Ile Ala Leu Ser Ala Ala Lys Tyr Thr
                325                 330                 335
Ala Ser Asn Pro Leu Trp Thr Val Thr Val Asn Leu Pro Ala Gly Thr
            340                 345                 350
Thr Val Leu Tyr Lys Phe Ile Asn Val Ala Ser Asp Gly Thr Val Thr
        355                 360                 365
Trp Gln Ala Asp Pro Asn Lys Ser Tyr Thr Val Pro Val Gly Cys Ala
    370                 375                 380
Thr Thr Ala Thr Val Ser Gly Ser Trp
385                 390

<210> 552
<211> 2019
<212> DNA
<213> Fungi
```

<400> 552
atggcttctt ctctgctttc ctccctcagc tccatcagca ctttcaacag cactcagatc 60
ctccaagctg ttcaacctat tatgcataca ttcgctcccc tccttttggt tggcagtctt 120
gccctccaaa ctatctatgc tttgcctggc cccagtaaag ttgagcaacg cgccgaaatc 180
ctaaaacgct cagtggattc ttttattgcc accgaaagtc ctattgcttt cagaaatctt 240
ttatgcaata tcggcgcaga tggagcttgt gcttcagggg cagcttcggg tattgtggtt 300
gcttcgcctg acaaagtgag cccagattat ttttatactt ggactcgaga ctccgcattg 360
gtattcaaag cccttgttga tacgtttatc agcagcagct attctgcatc acttcaacag 420
gaaattgaaa attacatcag ctcccaagct ggacttcaaa ccgtgtcaaa cccttcaggt 480
ggtttatcat ccggtggtct gggcgagccc aaatttaatg cggataagtc ggcattcaca 540
gggtcttggg gaagacctca acgagatgga cccgctttac gtgctactgc tatgattacc 600
tattcaaaat ggttgattgc caacggatac acctccactg tgcaaactat tgtttggccg 660
atcatcagaa atgatctgtc ttatgttact caatattgga accaaactgg ctttgatctc 720
tgggaagagg ttcaaggctc aagtttcttc acagtcgctg ctcaacatcg tgctcttgtt 780
gaaggtagtg cactggccgc tcaaattggt caaagctgta cctattgtga ttctcaagcc 840
ccacaggttc tttgtttcct ccaaagcttc tggagttcat cctcgggata catcatttcc 900
aacatcaacc aaaacagcgg aagaaacgga aaagatgcaa acagtatctt gagttctatc 960
cagacctttg atcctaccgc tgcatgtgat gccaccacgt tccaaccatg ctcagaccgt 1020
gctttggcaa atcacaaagt tgtgactgat tctttcagat ctatctacag catcaactct 1080
gggattgccg aaggtgtcgc cgtttctgtt ggacgctacc cagaagattc ttactatggt 1140
ggcaatcctt ggtatctcaa tacccttgct gctgccgaac agctttatga tgctttatac 1200
acctggaata gaattggctc tatcaccgta acttccactt ccctagcatt cttcaaagat 1260
ttcagctctt ccatcacagc cggaacatac gcatcctcca catcaaccta caccactctc 1320
tacaacgcca tcaaaaccta tgccgacggc tacgtcaatg tcgtcgcaac ctacgcacaa 1380
gccaacggtt ccctctcaga acaattcaac aaagccggcg gcgctcccct ctcagcctac 1440
gacctcacct ggtcctatgc tgccttcctc accgccgctg cccgtcgcgc aggtgtcgtc 1500
ccatactcat ggggtgagcc atccgcctcc agcgtcccag gaacctgttc tgctacctcg 1560
gccatcggaa cctattctac cgcaaccgcc acttcctggc ccgcatcgca aacccccagc 1620
ggaggaggag gaagtacccc cacaaccagc acatccgcca ccaccaagac aacttccacc 1680
tcgaccacca cttcttcatg cgcaacagcc acttcagtag ctgttacttt caatgaattg 1740
gtcactacct cctttggcca aacgatcaaa ctcgctggtt ccgtctctca attgggaagt 1800
tgggcaccgg cttctgcgat tgcattgagc gccgcgaagt acaccgctag taatccgttg 1860
tggaccgtga ctgttaacct accagctggc acgacggtat tgtataagtt tattaacgtg 1920
gcgagtgatg ggacggttac ctggcaggcg gatccaaata agagttatac tgttccggta 1980
ggctgtgcta cgactgctac tgttagtggt agctggtaa 2019

<210> 553
<211> 672
<212> PRT
<213> Fungi

<220>
<221> SIGNAL
<222> (1)...(23)

<221> DOMAIN
<222> (74)...(494)
<223> Glycosyl hydrolases family 15

<221> DOMAIN
<222> (572)...(668)
<223> Starch binding domain

<400> 553
Met Ala Ser Ser Leu Leu Ser Ser Leu Ser Ser Ile Ser Thr Phe Asn
 1   5   10   15
Ser Thr Gln Ile Leu Gln Ala Val Gln Pro Ile Met His Thr Phe Ala
            20   25   30
Pro Leu Leu Leu Val Gly Ser Leu Ala Leu Gln Thr Ile Tyr Ala Leu
        35   40   45
Pro Gly Pro Ser Lys Val Glu Gln Arg Ala Glu Ile Leu Lys Arg Ser
    50   55   60
Val Asp Ser Phe Ile Ala Thr Glu Ser Pro Ile Ala Phe Arg Asn Leu
65   70   75   80
Leu Cys Asn Ile Gly Ala Asp Gly Ala Cys Ala Ser Gly Ala Ala Ser
                85   90   95
Gly Ile Val Val Ala Ser Pro Asp Lys Val Ser Pro Asp Tyr Phe Tyr

```
            100                 105                 110
Thr Trp Thr Arg Asp Ser Ala Leu Val Phe Lys Ala Leu Val Asp Thr
        115                 120                 125
Phe Ile Ser Ser Ser Tyr Ser Ala Ser Leu Gln Gln Glu Ile Glu Asn
    130                 135                 140
Tyr Ile Ser Ser Gln Ala Gly Leu Gln Thr Val Ser Asn Pro Ser Gly
145                 150                 155                 160
Gly Leu Ser Ser Gly Gly Leu Gly Glu Pro Lys Phe Asn Ala Asp Lys
                165                 170                 175
Ser Ala Phe Thr Gly Ser Trp Gly Arg Pro Gln Arg Asp Gly Pro Ala
            180                 185                 190
Leu Arg Ala Thr Ala Met Ile Thr Tyr Ser Lys Trp Leu Ile Ala Asn
        195                 200                 205
Gly Tyr Thr Ser Thr Val Gln Thr Ile Val Trp Pro Ile Ile Arg Asn
    210                 215                 220
Asp Leu Ser Tyr Val Thr Gln Tyr Trp Asn Gln Thr Gly Phe Asp Leu
225                 230                 235                 240
Trp Glu Glu Val Gln Gly Ser Ser Phe Phe Thr Val Ala Ala Gln His
                245                 250                 255
Arg Ala Leu Val Glu Gly Ser Ala Leu Ala Ala Gln Ile Gly Gln Ser
            260                 265                 270
Cys Thr Tyr Cys Asp Ser Gln Ala Pro Gln Val Leu Cys Phe Leu Gln
        275                 280                 285
Ser Phe Trp Ser Ser Ser Ser Gly Tyr Ile Ile Ser Asn Ile Asn Gln
    290                 295                 300
Asn Ser Gly Arg Asn Gly Lys Asp Ala Asn Ser Ile Leu Ser Ser Ile
305                 310                 315                 320
Gln Thr Phe Asp Pro Thr Ala Ala Cys Asp Ala Thr Thr Phe Gln Pro
                325                 330                 335
Cys Ser Asp Arg Ala Leu Ala Asn His Lys Val Val Thr Asp Ser Phe
            340                 345                 350
Arg Ser Ile Tyr Ser Ile Asn Ser Gly Ile Ala Glu Gly Val Ala Val
        355                 360                 365
Ser Val Gly Arg Tyr Pro Glu Asp Ser Tyr Tyr Gly Gly Asn Pro Trp
    370                 375                 380
Tyr Leu Asn Thr Leu Ala Ala Ala Glu Gln Leu Tyr Asp Ala Leu Tyr
385                 390                 395                 400
Thr Trp Asn Arg Ile Gly Ser Ile Thr Val Thr Ser Thr Ser Leu Ala
                405                 410                 415
Phe Phe Lys Asp Phe Ser Ser Ser Ile Thr Ala Gly Thr Tyr Ala Ser
            420                 425                 430
Ser Thr Ser Thr Tyr Thr Thr Leu Tyr Asn Ala Ile Lys Thr Tyr Ala
        435                 440                 445
Asp Gly Tyr Val Asn Val Val Ala Thr Tyr Ala Gln Ala Asn Gly Ser
    450                 455                 460
Leu Ser Glu Gln Phe Asn Lys Ala Gly Gly Ala Pro Leu Ser Ala Tyr
465                 470                 475                 480
Asp Leu Thr Trp Ser Tyr Ala Ala Phe Leu Thr Ala Ala Ala Arg Arg
                485                 490                 495
Ala Gly Val Val Pro Tyr Ser Trp Gly Glu Pro Ser Ala Ser Ser Val
            500                 505                 510
Pro Gly Thr Cys Ser Ala Thr Ser Ala Ile Gly Thr Tyr Ser Thr Ala
        515                 520                 525
Thr Ala Thr Ser Trp Pro Ala Ser Gln Thr Pro Ser Gly Gly Gly Gly
    530                 535                 540
Ser Thr Pro Thr Thr Ser Thr Ser Ala Thr Thr Lys Thr Thr Ser Thr
545                 550                 555                 560
Ser Thr Thr Thr Ser Ser Cys Ala Thr Ala Thr Ser Val Ala Val Thr
                565                 570                 575
Phe Asn Glu Leu Val Thr Thr Ser Phe Gly Gln Thr Ile Lys Leu Ala
            580                 585                 590
Gly Ser Val Ser Gln Leu Gly Ser Trp Ala Pro Ala Ser Ala Ile Ala
        595                 600                 605
Leu Ser Ala Ala Lys Tyr Thr Ala Ser Asn Pro Leu Trp Thr Val Thr
    610                 615                 620
Val Asn Leu Pro Ala Gly Thr Thr Val Leu Tyr Lys Phe Ile Asn Val
625                 630                 635                 640
Ala Ser Asp Gly Thr Val Thr Trp Gln Ala Asp Pro Asn Lys Ser Tyr
```

645 650 655
Thr Val Pro Val Gly Cys Ala Thr Thr Ala Thr Val Ser Gly Ser Trp
660 665 670

<210> 554
<211> 1909
<212> DNA
<213> Cochliobolus heterostrophus ATCC 48331

<220>
<221> CDS
<222> (1)...(387)
<223> Exon

<221> CDS
<222> (451)...(990)
<223> Exon

<221> CDS
<222> (1040)...(1909)
<223> Exon

<400> 554
atg acg acg gca ttg tct tct ggc caa gtt gct ccc acc ccc cac acc 48
Met Thr Thr Ala Leu Ser Ser Gly Gln Val Ala Pro Thr Pro His Thr
1 5 10 15

gcc gcc gca atg aca atc act cct cga cca tgg tgg aag gat gct gtc 96
Ala Ala Ala Met Thr Ile Thr Pro Arg Pro Trp Trp Lys Asp Ala Val
20 25 30

gtc tac cag ata tac ccc gcc tcg ttc aag gac agc aac ggc gac ggc 144
Val Tyr Gln Ile Tyr Pro Ala Ser Phe Lys Asp Ser Asn Gly Asp Gly
35 40 45

att ggc gac ctc aat ggc atc atc tcg gaa ctc gac tac atc cgc tcc 192
Ile Gly Asp Leu Asn Gly Ile Ile Ser Glu Leu Asp Tyr Ile Arg Ser
50 55 60

att ggc gtc gat gtc atc tgg gtc tgc ccc atg tac gac tcg cct cag 240
Ile Gly Val Asp Val Ile Trp Val Cys Pro Met Tyr Asp Ser Pro Gln
65 70 75 80

gtc gac atg ggc tac gac atc cgc aac tac gag gat gtc tac aga cca 288
Val Asp Met Gly Tyr Asp Ile Arg Asn Tyr Glu Asp Val Tyr Arg Pro
85 90 95

tat gga acc gtc cag gac atg cag cgg ctg att gac gag acc cac tca 336
Tyr Gly Thr Val Gln Asp Met Gln Arg Leu Ile Asp Glu Thr His Ser
100 105 110

cgg ggc atg aag atc atc ctg gat ctc gtc gtc aac cac act tca gat 384
Arg Gly Met Lys Ile Ile Leu Asp Leu Val Val Asn His Thr Ser Asp
115 120 125

cag gtatgctcgc tctagtcacg tgatttgctg tggcattggc catcacatgc 437
Gln

tcactgccta cag cac caa tgg ttc caa gaa tca cgc tcg tca aag gac 486
His Gln Trp Phe Gln Glu Ser Arg Ser Ser Lys Asp
130 135 140

aat cca aag cgt gac tgg tac att tgg cgt ccc gcc cgc tac gtc gat 534
Asn Pro Lys Arg Asp Trp Tyr Ile Trp Arg Pro Ala Arg Tyr Val Asp
145 150 155

ggc gta cgc aaa ccc ccc aac aac tgg gtc tcc aac ttc acc ggc agt 582

```
Gly Val Arg Lys Pro Pro Asn Asn Trp Val Ser Asn Phe Thr Gly Ser
        160                 165                 170

gtc tgg gaa tgg gac gag cat acc caa gaa tac tac ctc cac ctc ttc      630
Val Trp Glu Trp Asp Glu His Thr Gln Glu Tyr Tyr Leu His Leu Phe
    175                 180                 185

tgt cct gag cag ccc gac ctc aac tgg gaa aat cca gag aca cgt aaa      678
Cys Pro Glu Gln Pro Asp Leu Asn Trp Glu Asn Pro Glu Thr Arg Lys
190                 195                 200                 205

gct ata tac gag tcg gcc atg gag ttc tgg ctc aag cgt ggt gtc gac      726
Ala Ile Tyr Glu Ser Ala Met Glu Phe Trp Leu Lys Arg Gly Val Asp
                210                 215                 220

ggc ttc cgc gtt gat acc gtc aac atg tac agt aag ggc gat atg cgt      774
Gly Phe Arg Val Asp Thr Val Asn Met Tyr Ser Lys Gly Asp Met Arg
            225                 230                 235

gac gca ccc atc acc gat ccg gga tca gag tgg cag ttt gcc ggc tac      822
Asp Ala Pro Ile Thr Asp Pro Gly Ser Glu Trp Gln Phe Ala Gly Tyr
        240                 245                 250

cag tac tgc aac ggc ccg cga atg gat gag ttc ctg ggc gag atg aat      870
Gln Tyr Cys Asn Gly Pro Arg Met Asp Glu Phe Leu Gly Glu Met Asn
    255                 260                 265

cag att ctg gaa aag tat gat gcc atg aca gtc ggc gag tgt cca cac      918
Gln Ile Leu Glu Lys Tyr Asp Ala Met Thr Val Gly Glu Cys Pro His
270                 275                 280                 285

aca ctt gat atc aac aga gta cac caa tac gta agc gcc aag gag aag      966
Thr Leu Asp Ile Asn Arg Val His Gln Tyr Val Ser Ala Lys Glu Lys
                290                 295                 300

cga ctc tcc atg gtg ttt caa ttc gtaagtggac ctggccctat tgcaataacc    1020
Arg Leu Ser Met Val Phe Gln Phe
            305

tagtgctcac ccaacccag gac gta gtc gat gtt ggg caa ggc ccc tac aaa     1072
                   Asp Val Val Asp Val Gly Gln Gly Pro Tyr Lys
                   310                 315                 320

ttc caa acc acg ccc aag aac tgg aca ctg cca caa ctc aag cgg gca     1120
Phe Gln Thr Thr Pro Lys Asn Trp Thr Leu Pro Gln Leu Lys Arg Ala
                325                 330                 335

ata gcc cgc act caa gat ctg att cgg cct cca tct gac agc tgg acc     1168
Ile Ala Arg Thr Gln Asp Leu Ile Arg Pro Pro Ser Asp Ser Trp Thr
            340                 345                 350

act gct ttt ctg gag aac cac gac caa gct cgg tca att acg cgc ttc     1216
Thr Ala Phe Leu Glu Asn His Asp Gln Ala Arg Ser Ile Thr Arg Phe
        355                 360                 365

act tct gac gcc cca cag cac cgt gtc gct ggc ggc aag atg ctt gcc     1264
Thr Ser Asp Ala Pro Gln His Arg Val Ala Gly Gly Lys Met Leu Ala
    370                 375                 380

ctc atg ctg agc gct ctt agc ggc acg ctg ttc atc tac caa ggc cag     1312
Leu Met Leu Ser Ala Leu Ser Gly Thr Leu Phe Ile Tyr Gln Gly Gln
385                 390                 395                 400

gaa atc ggc atg aca aac ttc ccc gaa tcg tgg gac atg agc gaa tac     1360
Glu Ile Gly Met Thr Asn Phe Pro Glu Ser Trp Asp Met Ser Glu Tyr
                405                 410                 415

aag gac gtc gag tca tcc aat tac tac aag atg gtt gcc aag cgc acc     1408
```

```
Lys Asp Val Glu Ser Ser Asn Tyr Tyr Lys Met Val Ala Lys Arg Thr
            420                 425                 430

aac aac gat gtc gat gcg ctt gct gct gca cac aag tcg ttg cag cac     1456
Asn Asn Asp Val Asp Ala Leu Ala Ala Ala His Lys Ser Leu Gln His
        435                 440                 445

ctt gcg cgc gac cac tcg cgc gtg ccc atg agc tgg agt acc gca ccg     1504
Leu Ala Arg Asp His Ser Arg Val Pro Met Ser Trp Ser Thr Ala Pro
    450                 455                 460

tac aat ggc ttc agt ccg ccc gat gca aaa gac aag cca tgg atg cgg     1552
Tyr Asn Gly Phe Ser Pro Pro Asp Ala Lys Asp Lys Pro Trp Met Arg
465                 470                 475                 480

ccg cta gaa gac gca gac att tgc aat gca aag gcg caa caa aat gac     1600
Pro Leu Glu Asp Ala Asp Ile Cys Asn Ala Lys Ala Gln Gln Asn Asp
                485                 490                 495

aag aca tct gtg ctg ggt ttc tgg aaa cac atg ctg cag ctg cgc aag     1648
Lys Thr Ser Val Leu Gly Phe Trp Lys His Met Leu Gln Leu Arg Lys
            500                 505                 510

gag cac aag gat ttg ctc gtc ttc ggt cag tac gac gat ctt gac gtt     1696
Glu His Lys Asp Leu Leu Val Phe Gly Gln Tyr Asp Asp Leu Asp Val
        515                 520                 525

gac aat gag caa ttt tac att ttc agc aag acg tgg cag gga aag cgg     1744
Asp Asn Glu Gln Phe Tyr Ile Phe Ser Lys Thr Trp Gln Gly Lys Arg
    530                 535                 540

gcg ctg tgt att tgc aac ttt aca gat gag agc aag cag ctg gtt ttg     1792
Ala Leu Cys Ile Cys Asn Phe Thr Asp Glu Ser Lys Gln Leu Val Leu
545                 550                 555                 560

cca gag agc gtg gca gcg cag aaa atg cag ctg ctg gta agt agt cgg     1840
Pro Glu Ser Val Ala Ala Gln Lys Met Gln Leu Leu Val Ser Ser Arg
                565                 570                 575

caa gac gag ggc gtc gag gaa aag acg cta gcg ccg tac gag gga agg     1888
Gln Asp Glu Gly Val Glu Glu Lys Thr Leu Ala Pro Tyr Glu Gly Arg
            580                 585                 590

gtc tat ctt tat gtg caa tga                                       1909
Val Tyr Leu Tyr Val Gln  *
        595


<210> 555
<211> 129
<212> PRT
<213> Cochliobolus heterostrophus ATCC 48331

<400> 555
Met Thr Thr Ala Leu Ser Ser Gly Gln Val Ala Pro Thr Pro His Thr
 1               5                  10                  15
Ala Ala Ala Met Thr Ile Thr Pro Arg Pro Trp Trp Lys Asp Ala Val
            20                  25                  30
Val Tyr Gln Ile Tyr Pro Ala Ser Phe Lys Asp Ser Asn Gly Asp Gly
        35                  40                  45
Ile Gly Asp Leu Asn Gly Ile Ile Ser Glu Leu Asp Tyr Ile Arg Ser
    50                  55                  60
Ile Gly Val Asp Val Ile Trp Val Cys Pro Met Tyr Asp Ser Pro Gln
65                  70                  75                  80
Val Asp Met Gly Tyr Asp Ile Arg Asn Tyr Glu Asp Val Tyr Arg Pro
                85                  90                  95
Tyr Gly Thr Val Gln Asp Met Gln Arg Leu Ile Asp Glu Thr His Ser
            100                 105                 110
```

Arg Gly Met Lys Ile Ile Leu Asp Leu Val Val Asn His Thr Ser Asp
115 120 125
Gln

<210> 556
<211> 180
<212> PRT
<213> Cochliobolus heterostrophus ATCC 48331

<400> 556
His Gln Trp Phe Gln Glu Ser Arg Ser Ser Lys Asp Asn Pro Lys Arg
1 5 10 15
Asp Trp Tyr Ile Trp Arg Pro Ala Arg Tyr Val Asp Gly Val Arg Lys
20 25 30
Pro Pro Asn Asn Trp Val Ser Asn Phe Thr Gly Ser Val Trp Glu Trp
35 40 45
Asp Glu His Thr Gln Glu Tyr Tyr Leu His Leu Phe Cys Pro Glu Gln
50 55 60
Pro Asp Leu Asn Trp Glu Asn Pro Glu Thr Arg Lys Ala Ile Tyr Glu
65 70 75 80
Ser Ala Met Glu Phe Trp Leu Lys Arg Gly Val Asp Gly Phe Arg Val
85 90 95
Asp Thr Val Asn Met Tyr Ser Lys Gly Asp Met Arg Asp Ala Pro Ile
100 105 110
Thr Asp Pro Gly Ser Glu Trp Gln Phe Ala Gly Tyr Gln Tyr Cys Asn
115 120 125
Gly Pro Arg Met Asp Glu Phe Leu Gly Glu Met Asn Gln Ile Leu Glu
130 135 140
Lys Tyr Asp Ala Met Thr Val Gly Glu Cys Pro His Thr Leu Asp Ile
145 150 155 160
Asn Arg Val His Gln Tyr Val Ser Ala Lys Glu Lys Arg Leu Ser Met
165 170 175
Val Phe Gln Phe
180

<210> 557
<211> 289
<212> PRT
<213> Cochliobolus heterostrophus ATCC 48331

<400> 557
Asp Val Val Asp Val Gly Gln Gly Pro Tyr Lys Phe Gln Thr Thr Pro
1 5 10 15
Lys Asn Trp Thr Leu Pro Gln Leu Lys Arg Ala Ile Ala Arg Thr Gln
20 25 30
Asp Leu Ile Arg Pro Pro Ser Asp Ser Trp Thr Thr Ala Phe Leu Glu
35 40 45
Asn His Asp Gln Ala Arg Ser Ile Thr Arg Phe Thr Ser Asp Ala Pro
50 55 60
Gln His Arg Val Ala Gly Gly Lys Met Leu Ala Leu Met Leu Ser Ala
65 70 75 80
Leu Ser Gly Thr Leu Phe Ile Tyr Gln Gly Gln Glu Ile Gly Met Thr
85 90 95
Asn Phe Pro Glu Ser Trp Asp Met Ser Glu Tyr Lys Asp Val Glu Ser
100 105 110
Ser Asn Tyr Tyr Lys Met Val Ala Lys Arg Thr Asn Asn Asp Val Asp
115 120 125
Ala Leu Ala Ala Ala His Lys Ser Leu Gln His Leu Ala Arg Asp His
130 135 140
Ser Arg Val Pro Met Ser Trp Ser Thr Ala Pro Tyr Asn Gly Phe Ser
145 150 155 160
Pro Pro Asp Ala Lys Asp Lys Pro Trp Met Arg Pro Leu Glu Asp Ala
165 170 175
Asp Ile Cys Asn Ala Lys Ala Gln Gln Asn Asp Lys Thr Ser Val Leu
180 185 190
Gly Phe Trp Lys His Met Leu Gln Leu Arg Lys Glu His Lys Asp Leu
195 200 205

```
Leu Val Phe Gly Gln Tyr Asp Asp Leu Asp Val Asp Asn Glu Gln Phe
    210                 215                 220
Tyr Ile Phe Ser Lys Thr Trp Gln Gly Lys Arg Ala Leu Cys Ile Cys
225                 230                 235                 240
Asn Phe Thr Asp Glu Ser Lys Gln Leu Val Leu Pro Glu Ser Val Ala
                245                 250                 255
Ala Gln Lys Met Gln Leu Leu Val Ser Ser Arg Gln Asp Glu Gly Val
            260                 265                 270
Glu Glu Lys Thr Leu Ala Pro Tyr Glu Gly Arg Val Tyr Leu Tyr Val
        275                 280                 285
Gln
```

<210> 558
<211> 1797
<212> DNA
<213> Cochliobolus heterostrophus ATCC 48331

<400> 558

```
atgacgacgg cattgtcttc tggccaagtt gctcccaccc cccacaccgc cgccgcaatg      60
acaatcactc ctcgaccatg gtggaaggat gctgtcgtct accagatata ccccgcctcg     120
ttcaaggaca gcaacggcga cggcattggc gacctcaatg gcatcatctc ggaactcgac     180
tacatccgct ccattggcgt cgatgtcatc tgggtctgcc ccatgtacga ctcgcctcag     240
gtcgacatgg gctacgacat ccgcaactac gaggatgtct acagaccata tggaaccgtc     300
caggacatgc agcggctgat tgacgagacc cactcacggg gcatgaagat catcctggat     360
ctcgtcgtca accacacttc agatcagcac caatggttcc aagaatcacg ctcgtcaaag     420
gacaatccaa agcgtgactg gtacatttgg cgtcccgccc gctacgtcga tggcgtacgc     480
aaaccccccа acaactgggt ctccaacttc accggcagtg tctgggaatg ggacgagcat     540
acccaagaat actacctcca cctcttctgt cctgagcagc ccgacctcaa ctgggaaaat     600
ccagagacac gtaaagctat atacgagtcg gccatggagt tctggctcaa gcgtggtgtc     660
gacggcttcc gcgttgatac cgtcaacatg tacagtaagg gcgatatgcg tgacgcaccc     720
atcaccgatc cgggatcaga gtggcagttt gccggctacc agtactgcaa cggcccgcga     780
atggatgagt tcctgggcga gatgaatcag attctggaaa agtatgatgc catgacagtc     840
ggcgagtgtc cacacacact tgatatcaac agagtacacc aatacgtaag cgccaaggag     900
aagcgactct ccatggtgtt tcaattcgac gtagtcgatg ttgggcaagg cccctacaaa     960
ttccaaacca cgcccaagaa ctggacactg ccacaactca agcgggcaat agcccgcact    1020
caagatctga ttcggcctcc atctgacagc tggaccactg cttttctgga gaaccacgac    1080
caagctcggt caattacgcg cttcacttct gacgccccac agcaccgtgt cgctggcggc    1140
aagatgcttg ccctcatgct gagcgctctt agcggcacgc tgttcatcta ccaaggccag    1200
gaaatcggca tgacaaactt ccccgaatcg tgggacatga gcgaatacaa ggacgtcgag    1260
tcatccaatt actacaagat ggttgccaag cgcaccaaca acgatgtcga tgcgcttgct    1320
gctgcacaca agtcgttgca gcaccttgcg cgcgaccact cgcgcgtgcc catgagctgg    1380
agtaccgcac cgtacaatgg cttcagtccg cccgatgcaa aagacaagcc atggatgcgg    1440
ccgctagaag acgcagacat ttgcaatgca aaggcgcaac aaaatgacaa gacatctgtg    1500
ctgggtttct ggaaacacat gctgcagctg cgcaaggagc acaaggattt gctcgtcttc    1560
ggtcagtacg acgatcttga cgttgacaat gagcaatttt acattttcag caagacgtgg    1620
cagggaaagc gggcgctgtg tatttgcaac tttacagatg agagcaagca gctggttttg    1680
ccagagagcg tggcagcgca gaaaatgcag ctgctggtaa gtagtcggca agacgagggc    1740
gtcgaggaaa agacgctagc gccgtacgag ggaagggtct atctttatgt gcaatga       1797
```

<210> 559
<211> 598
<212> PRT
<213> Cochliobolus heterostrophus ATCC 48331

<220>
<221> SIGNAL
<222> (1)...(19)

<221> DOMAIN
<222> (35)...(443)
<223> Alpha amylase, catalytic domain

<400> 559

```
Met Thr Thr Ala Leu Ser Ser Gly Gln Val Ala Pro Thr Pro His Thr
 1               5                  10                  15
Ala Ala Ala Met Thr Ile Thr Pro Arg Pro Trp Trp Lys Asp Ala Val
            20                  25                  30
```

```
Val Tyr Gln Ile Tyr Pro Ala Ser Phe Lys Asp Ser Asn Gly Asp Gly
        35                  40                  45
Ile Gly Asp Leu Asn Gly Ile Ile Ser Glu Leu Asp Tyr Ile Arg Ser
    50                  55                  60
Ile Gly Val Asp Val Ile Trp Val Cys Pro Met Tyr Asp Ser Pro Gln
65                  70                  75                  80
Val Asp Met Gly Tyr Asp Ile Arg Asn Tyr Glu Asp Val Tyr Arg Pro
                85                  90                  95
Tyr Gly Thr Val Gln Asp Met Gln Arg Leu Ile Asp Glu Thr His Ser
            100                 105                 110
Arg Gly Met Lys Ile Ile Leu Asp Leu Val Val Asn His Thr Ser Asp
        115                 120                 125
Gln His Gln Trp Phe Gln Glu Ser Arg Ser Ser Lys Asp Asn Pro Lys
    130                 135                 140
Arg Asp Trp Tyr Ile Trp Arg Pro Ala Arg Tyr Val Asp Gly Val Arg
145                 150                 155                 160
Lys Pro Pro Asn Asn Trp Val Ser Asn Phe Thr Gly Ser Val Trp Glu
                165                 170                 175
Trp Asp Glu His Thr Gln Glu Tyr Tyr Leu His Leu Phe Cys Pro Glu
            180                 185                 190
Gln Pro Asp Leu Asn Trp Glu Asn Pro Glu Thr Arg Lys Ala Ile Tyr
        195                 200                 205
Glu Ser Ala Met Glu Phe Trp Leu Lys Arg Gly Val Asp Gly Phe Arg
    210                 215                 220
Val Asp Thr Val Asn Met Tyr Ser Lys Gly Asp Met Arg Asp Ala Pro
225                 230                 235                 240
Ile Thr Asp Pro Gly Ser Glu Trp Gln Phe Ala Gly Tyr Gln Tyr Cys
                245                 250                 255
Asn Gly Pro Arg Met Asp Glu Phe Leu Gly Glu Met Asn Gln Ile Leu
            260                 265                 270
Glu Lys Tyr Asp Ala Met Thr Val Gly Glu Cys Pro His Thr Leu Asp
        275                 280                 285
Ile Asn Arg Val His Gln Tyr Val Ser Ala Lys Glu Lys Arg Leu Ser
    290                 295                 300
Met Val Phe Gln Phe Asp Val Val Asp Val Gly Gln Gly Pro Tyr Lys
305                 310                 315                 320
Phe Gln Thr Thr Pro Lys Asn Trp Thr Leu Pro Gln Leu Lys Arg Ala
                325                 330                 335
Ile Ala Arg Thr Gln Asp Leu Ile Arg Pro Pro Ser Asp Ser Trp Thr
            340                 345                 350
Thr Ala Phe Leu Glu Asn His Asp Gln Ala Arg Ser Ile Thr Arg Phe
        355                 360                 365
Thr Ser Asp Ala Pro Gln His Arg Val Ala Gly Gly Lys Met Leu Ala
    370                 375                 380
Leu Met Leu Ser Ala Leu Ser Gly Thr Leu Phe Ile Tyr Gln Gly Gln
385                 390                 395                 400
Glu Ile Gly Met Thr Asn Phe Pro Glu Ser Trp Asp Met Ser Glu Tyr
                405                 410                 415
Lys Asp Val Glu Ser Ser Asn Tyr Tyr Lys Met Val Ala Lys Arg Thr
            420                 425                 430
Asn Asn Asp Val Asp Ala Leu Ala Ala Ala His Lys Ser Leu Gln His
        435                 440                 445
Leu Ala Arg Asp His Ser Arg Val Pro Met Ser Trp Ser Thr Ala Pro
    450                 455                 460
Tyr Asn Gly Phe Ser Pro Pro Asp Ala Lys Asp Lys Pro Trp Met Arg
465                 470                 475                 480
Pro Leu Glu Asp Ala Asp Ile Cys Asn Ala Lys Ala Gln Gln Asn Asp
                485                 490                 495
Lys Thr Ser Val Leu Gly Phe Trp Lys His Met Leu Gln Leu Arg Lys
            500                 505                 510
Glu His Lys Asp Leu Leu Val Phe Gly Gln Tyr Asp Asp Leu Asp Val
        515                 520                 525
Asp Asn Glu Gln Phe Tyr Ile Phe Ser Lys Thr Trp Gln Gly Lys Arg
    530                 535                 540
Ala Leu Cys Ile Cys Asn Phe Thr Asp Glu Ser Lys Gln Leu Val Leu
545                 550                 555                 560
Pro Glu Ser Val Ala Ala Gln Lys Met Gln Leu Leu Val Ser Ser Arg
                565                 570                 575
```

```
Gln Asp Glu Gly Val Glu Glu Lys Thr Leu Ala Pro Tyr Glu Gly Arg
            580                 585                 590
Val Tyr Leu Tyr Val Gln
        595

<210> 560
<211> 1897
<212> DNA
<213> Fungi

<220>
<221> CDS
<222> (1)...(276)
<223> Exon

<221> CDS
<222> (443)...(577)
<223> Exon

<221> CDS
<222> (648)...(909)
<223> Exon

<221> CDS
<222> (1047)...(1897)
<223> Exon

<400> 560
atg ttg acc acg tcg gag agg aaa act tcg acg gca ttc gtt act tgg       48
Met Leu Thr Thr Ser Glu Arg Lys Thr Ser Thr Ala Phe Val Thr Trp
 1               5                  10                  15

tca atg ttg tgg gtg gtg ttg ttg aca tcc ttc gtg aaa gat gtc cac       96
Ser Met Leu Trp Val Val Leu Leu Thr Ser Phe Val Lys Asp Val His
             20                  25                  30

gct gcc act tca gac gat tgg aga agc caa tct ata tat caa att atc      144
Ala Ala Thr Ser Asp Asp Trp Arg Ser Gln Ser Ile Tyr Gln Ile Ile
         35                  40                  45

aca gat cga ttc gcc aga acg gat ggc tcg act acc tat gcg tgt aac      192
Thr Asp Arg Phe Ala Arg Thr Asp Gly Ser Thr Thr Tyr Ala Cys Asn
     50                  55                  60

aac aca gca cag ctg tat tgc ggt gga act ttt caa gga att atc aac      240
Asn Thr Ala Gln Leu Tyr Cys Gly Gly Thr Phe Gln Gly Ile Ile Asn
 65                  70                  75                  80

cag cta gat tac atc cag gat atg gga ttt acg gct gtgagtttgt           286
Gln Leu Asp Tyr Ile Gln Asp Met Gly Phe Thr Ala
                 85                  90

ctatctcagc ttctccttac gctcttcgca cgaacgatcc ttctctctgt tgtgtctcgt    346
acgaccgtcg cgacgaaaga tgggaaaatc caccagccta gacatcgtga actgacttaa    406
tgccaaactt agatctggat ttcaccggtg gtacag aat atc aac caa aca acg      460
                                       Asn Ile Asn Gln Thr Thr
                                                   95

gca tat ggg caa ggc tat cat gga ttt tgg agt caa gat atc act aag      508
Ala Tyr Gly Gln Gly Tyr His Gly Phe Trp Ser Gln Asp Ile Thr Lys
    100                 105                 110

atc aac gaa cat ttc gga aca gca gac gat ctc aaa ctc ttg tcg tca      556
Ile Asn Glu His Phe Gly Thr Ala Asp Asp Leu Lys Leu Leu Ser Ser
115                 120                 125                 130

acc ttg cac gat cgt ggc atg gtaagctgca ttcacattcg caacccaatc         607
Thr Leu His Asp Arg Gly Met
```

```
                135

acatggaaca taagctctaa ctcaactgtg cgcattatag tat cta atg atc gat      662
                                                Tyr Leu Met Ile Asp
                                                        140

gcg gtc att aac gac atg gca tac gca ggg aaa ggg agc gca gtc gat      710
Ala Val Ile Asn Asp Met Ala Tyr Ala Gly Lys Gly Ser Ala Val Asp
        145                 150                 155

tac tcc act cta gtt ccg ttc aac aaa aaa gaa tac ttc cat cct ttc      758
Tyr Ser Thr Leu Val Pro Phe Asn Lys Lys Glu Tyr Phe His Pro Phe
    160                 165                 170

tgt tac att acc gat tat tcc aac gca aca aac ttt caa aac tgt tgg      806
Cys Tyr Ile Thr Asp Tyr Ser Asn Ala Thr Asn Phe Gln Asn Cys Trp
175                 180                 185                 190

ctg gga gac gat aca gtt tca ttg cca gat ctg gat acc gaa tct gat      854
Leu Gly Asp Asp Thr Val Ser Leu Pro Asp Leu Asp Thr Glu Ser Asp
                195                 200                 205

ttt gtc aag caa aca tgg gaa acc tgg gta aca gaa atg gtg gcc aat      902
Phe Val Lys Gln Thr Trp Glu Thr Trp Val Thr Glu Met Val Ala Asn
            210                 215                 220

tat tct c gtatgtactt ctgatcatga atctcactca ccccttgctc tttctggaag     959
Tyr Ser


tcttggctgt ggggaaaatc tggggaatat taatgcacaa tacattccat tgcaacaaga   1019
gtgggctaac ataatctcga actttag tt gat ggt ctt cgt atc gat gcc gcc    1072
                              Leu Asp Gly Leu Arg Ile Asp Ala Ala
                              225                 230

aaa cac gtc gac aag cct ttc tgg cct ggc ttc caa cag gcc gct aat     1120
Lys His Val Asp Lys Pro Phe Trp Pro Gly Phe Gln Gln Ala Ala Asn
    235                 240                 245

aca ttc act att ggt gaa gtt ttt gac gag ctg cct acc aat gca tgc     1168
Thr Phe Thr Ile Gly Glu Val Phe Asp Glu Leu Pro Thr Asn Ala Cys
250                 255                 260                 265

gaa tgg gcc gtc gat gct ctc tct tct gtc cta aat tat cct gcc tgg     1216
Glu Trp Ala Val Asp Ala Leu Ser Ser Val Leu Asn Tyr Pro Ala Trp
                270                 275                 280

tat tac att acg tct atc ctc tca aac tca acg aat ggt atg gga agt     1264
Tyr Tyr Ile Thr Ser Ile Leu Ser Asn Ser Thr Asn Gly Met Gly Ser
            285                 290                 295

tta agc tac caa ttc gat caa acc cag caa tac tgt tac gat act aca     1312
Leu Ser Tyr Gln Phe Asp Gln Thr Gln Gln Tyr Cys Tyr Asp Thr Thr
        300                 305                 310

ctt ctt ggt act ttc agt gag aat cat gat gtt gct cgt ttt gga tcc     1360
Leu Leu Gly Thr Phe Ser Glu Asn His Asp Val Ala Arg Phe Gly Ser
    315                 320                 325

tac acg agt gac gtg tct caa agg aaa aat gcc ctc acg ttc gac ttt     1408
Tyr Thr Ser Asp Val Ser Gln Arg Lys Asn Ala Leu Thr Phe Asp Phe
330                 335                 340                 345

ttt aca gac ggc att cct att gta tat tac gga gcc gag caa ggc ctt     1456
Phe Thr Asp Gly Ile Pro Ile Val Tyr Tyr Gly Ala Glu Gln Gly Leu
                350                 355                 360

gac ggt gaa aac gat cct caa aac cgt ggg gct ctt tgg cta aat cca     1504
```

```
Asp Gly Glu Asn Asp Pro Gln Asn Arg Gly Ala Leu Trp Leu Asn Pro
            365                 370                 375

acc ggg tat gat act acc gca gtt ctg tat caa cat atc aaa acc atg     1552
Thr Gly Tyr Asp Thr Thr Ala Val Leu Tyr Gln His Ile Lys Thr Met
        380                 385                 390

aat act gcc cgc aat gcc gtc aac aat tac atg ata gca acc aat tac     1600
Asn Thr Ala Arg Asn Ala Val Asn Asn Tyr Met Ile Ala Thr Asn Tyr
    395                 400                 405

tct aat tgg agt cct tac tgg gca tac aaa gca tcg gtc ata aag caa     1648
Ser Asn Trp Ser Pro Tyr Trp Ala Tyr Lys Ala Ser Val Ile Lys Gln
410                 415                 420                 425

gca gac gac gta ttg gtg ttt agg aaa ggg ttg gtc cat tcg atc gtc     1696
Ala Asp Asp Val Leu Val Phe Arg Lys Gly Leu Val His Ser Ile Val
                430                 435                 440

aca gct atc acc aat gtc ggt acc gag gga gcg act gtt ggg cct tac     1744
Thr Ala Ile Thr Asn Val Gly Thr Glu Gly Ala Thr Val Gly Pro Tyr
            445                 450                 455

tac atc gaa gac acc aat ttc agc gaa gga aat ctc ctc att gag atc     1792
Tyr Ile Glu Asp Thr Asn Phe Ser Glu Gly Asn Leu Leu Ile Glu Ile
        460                 465                 470

ttg agc tgt aac agc act gtc gcg gga gtc ggt gga gct ttc acc ctc     1840
Leu Ser Cys Asn Ser Thr Val Ala Gly Val Gly Gly Ala Phe Thr Leu
    475                 480                 485

acc ttg act agc gga gag cca cag gta tgt cta tat agt ttt aac gca     1888
Thr Leu Thr Ser Gly Glu Pro Gln Val Cys Leu Tyr Ser Phe Asn Ala
490                 495                 500                 505

tat gcg taa                                                        1897
Tyr Ala  *



<210> 561
<211> 92
<212> PRT
<213> Fungi

<400> 561
Met Leu Thr Thr Ser Glu Arg Lys Thr Ser Thr Ala Phe Val Thr Trp
 1               5                  10                  15
Ser Met Leu Trp Val Val Leu Leu Thr Ser Phe Val Lys Asp Val His
            20                  25                  30
Ala Ala Thr Ser Asp Asp Trp Arg Ser Gln Ser Ile Tyr Gln Ile Ile
        35                  40                  45
Thr Asp Arg Phe Ala Arg Thr Asp Gly Ser Thr Thr Tyr Ala Cys Asn
    50                  55                  60
Asn Thr Ala Gln Leu Tyr Cys Gly Gly Thr Phe Gln Gly Ile Ile Asn
65                  70                  75                  80
Gln Leu Asp Tyr Ile Gln Asp Met Gly Phe Thr Ala
                85                  90

<210> 562
<211> 45
<212> PRT
<213> Fungi

<400> 562
Asn Ile Asn Gln Thr Thr Ala Tyr Gly Gln Gly Tyr His Gly Phe Trp
 1               5                  10                  15
Ser Gln Asp Ile Thr Lys Ile Asn Glu His Phe Gly Thr Ala Asp Asp
```

20 25 30
Leu Lys Leu Leu Ser Ser Thr Leu His Asp Arg Gly Met
35 40 45

<210> 563
<211> 87
<212> PRT
<213> Fungi

<400> 563
Tyr Leu Met Ile Asp Ala Val Ile Asn Asp Met Ala Tyr Ala Gly Lys
1 5 10 15
Gly Ser Ala Val Asp Tyr Ser Thr Leu Val Pro Phe Asn Lys Lys Glu
20 25 30
Tyr Phe His Pro Phe Cys Tyr Ile Thr Asp Tyr Ser Asn Ala Thr Asn
35 40 45
Phe Gln Asn Cys Trp Leu Gly Asp Asp Thr Val Ser Leu Pro Asp Leu
50 55 60
Asp Thr Glu Ser Asp Phe Val Lys Gln Thr Trp Glu Thr Trp Val Thr
65 70 75 80
Glu Met Val Ala Asn Tyr Ser
85

<210> 564
<211> 283
<212> PRT
<213> Fungi

<400> 564
Leu Asp Gly Leu Arg Ile Asp Ala Ala Lys His Val Asp Lys Pro Phe
1 5 10 15
Trp Pro Gly Phe Gln Gln Ala Ala Asn Thr Phe Thr Ile Gly Glu Val
20 25 30
Phe Asp Glu Leu Pro Thr Asn Ala Cys Glu Trp Ala Val Asp Ala Leu
35 40 45
Ser Ser Val Leu Asn Tyr Pro Ala Trp Tyr Tyr Ile Thr Ser Ile Leu
50 55 60
Ser Asn Ser Thr Asn Gly Met Gly Ser Leu Ser Tyr Gln Phe Asp Gln
65 70 75 80
Thr Gln Gln Tyr Cys Tyr Asp Thr Thr Leu Leu Gly Thr Phe Ser Glu
85 90 95
Asn His Asp Val Ala Arg Phe Gly Ser Tyr Thr Ser Asp Val Ser Gln
100 105 110
Arg Lys Asn Ala Leu Thr Phe Asp Phe Phe Thr Asp Gly Ile Pro Ile
115 120 125
Val Tyr Tyr Gly Ala Glu Gln Gly Leu Asp Gly Glu Asn Asp Pro Gln
130 135 140
Asn Arg Gly Ala Leu Trp Leu Asn Pro Thr Gly Tyr Asp Thr Thr Ala
145 150 155 160
Val Leu Tyr Gln His Ile Lys Thr Met Asn Thr Ala Arg Asn Ala Val
165 170 175
Asn Asn Tyr Met Ile Ala Thr Asn Tyr Ser Asn Trp Ser Pro Tyr Trp
180 185 190
Ala Tyr Lys Ala Ser Val Ile Lys Gln Ala Asp Asp Val Leu Val Phe
195 200 205
Arg Lys Gly Leu Val His Ser Ile Val Thr Ala Ile Thr Asn Val Gly
210 215 220
Thr Glu Gly Ala Thr Val Gly Pro Tyr Tyr Ile Glu Asp Thr Asn Phe
225 230 235 240
Ser Glu Gly Asn Leu Leu Ile Glu Ile Leu Ser Cys Asn Ser Thr Val
245 250 255
Ala Gly Val Gly Gly Ala Phe Thr Leu Thr Leu Thr Ser Gly Glu Pro
260 265 270
Gln Val Cys Leu Tyr Ser Phe Asn Ala Tyr Ala
275 280

<210> 565
<211> 1524

<212> DNA
<213> Fungi

<400> 565
atgttgacca cgtcggagag gaaaacttcg acggcattcg ttacttggtc aatgttgtgg 60
gtggtgttgt tgacatcctt cgtgaaagat gtccacgctg ccacttcaga cgattggaga 120
agccaatcta tatatcaaat tatcacagat cgattcgcca gaacggatgg ctcgactacc 180
tatgcgtgta acaacacagc acagctgtat tgcggtggaa cttttcaagg aattatcaac 240
cagctagatt acatccagga tatgggattt acggctaata tcaaccaaac aacggcatat 300
gggcaaggct atcatggatt ttggagtcaa gatatcacta agatcaacga acatttcgga 360
acagcagacg atctcaaact cttgtcgtca accttgcacg atcgtggcat gtatctaatg 420
atcgatgcgg tcattaacga catggcatac gcagggaaag ggagcgcagt cgattactcc 480
actctagttc cgttcaacaa aaaagaatac ttccatcctt tctgttacat taccgattat 540
tccaacgcaa caaactttca aaactgttgg ctgggagacg atacagtttc attgccagat 600
ctggataccg aatctgattt tgtcaagcaa acatgggaaa cctgggtaac agaaatggtg 660
gccaattatt ctcttgatgg tcttcgtatc gatgccgcca aacacgtcga caagcctttc 720
tggcctggct tccaacaggc cgctaataca ttcactattg gtgaagtttt tgacgagctg 780
cctaccaatg catgcgaatg ggccgtcgat gctctctctt ctgtcctaaa ttatcctgcc 840
tggtattaca ttacgtctat cctctcaaac tcaacgaatg gtatgggaag tttaagctac 900
caattcgatc aaacccagca atactgttac gatactacac ttcttggtac tttcagtgag 960
aatcatgatg ttgctcgttt tggatcctac acgagtgacg tgtctcaaag gaaaaatgcc 1020
ctcacgttcg actttttac agacggcatt cctattgtat attacggagc cgagcaaggc 1080
cttgacggtg aaaacgatcc tcaaaaccgt ggggctcttt ggctaaatcc aaccgggtat 1140
gatactaccg cagttctgta tcaacatatc aaaaccatga atactgcccg caatgccgtc 1200
aacaattaca tgatagcaac caattactct aattggagtc cttactgggc atacaaagca 1260
tcggtcataa agcaagcaga cgacgtattg gtgtttagga aagggttggt ccattcgatc 1320
gtcacagcta tcaccaatgt cggtaccgag ggagcgactg ttgggcctta ctacatcgaa 1380
gacaccaatt tcagcgaagg aaatctcctc attgagatct tgagctgtaa cagcactgtc 1440
gcgggagtcg gtggagcttt caccctcacc ttgactagcg gagagccaca ggtatgtcta 1500
tatagtttta acgcatatgc gtaa 1524

<210> 566
<211> 507
<212> PRT
<213> Fungi

<220>
<221> SIGNAL
<222> (1)...(33)

<221> DOMAIN
<222> (46)...(397)
<223> Alpha amylase, catalytic domain

<400> 566
Met Leu Thr Thr Ser Glu Arg Lys Thr Ser Thr Ala Phe Val Thr Trp
 1   5   10   15
Ser Met Leu Trp Val Val Leu Leu Thr Ser Phe Val Lys Asp Val His
            20   25   30
Ala Ala Thr Ser Asp Asp Trp Arg Ser Gln Ser Ile Tyr Gln Ile Ile
        35   40   45
Thr Asp Arg Phe Ala Arg Thr Asp Gly Ser Thr Thr Tyr Ala Cys Asn
    50   55   60
Asn Thr Ala Gln Leu Tyr Cys Gly Gly Thr Phe Gln Gly Ile Ile Asn
65   70   75   80
Gln Leu Asp Tyr Ile Gln Asp Met Gly Phe Thr Ala Asn Ile Asn Gln
                85   90   95
Thr Thr Ala Tyr Gly Gln Gly Tyr His Gly Phe Trp Ser Gln Asp Ile
            100   105   110
Thr Lys Ile Asn Glu His Phe Gly Thr Ala Asp Asp Leu Lys Leu Leu
        115   120   125
Ser Ser Thr Leu His Asp Arg Gly Met Tyr Leu Met Ile Asp Ala Val
    130   135   140
Ile Asn Asp Met Ala Tyr Ala Gly Lys Gly Ser Ala Val Asp Tyr Ser
145   150   155   160
Thr Leu Val Pro Phe Asn Lys Lys Glu Tyr Phe His Pro Phe Cys Tyr
                165   170   175
Ile Thr Asp Tyr Ser Asn Ala Thr Asn Phe Gln Asn Cys Trp Leu Gly

```
            180                 185                 190
Asp Asp Thr Val Ser Leu Pro Asp Leu Asp Thr Glu Ser Asp Phe Val
        195                 200                 205
Lys Gln Thr Trp Glu Thr Trp Val Thr Glu Met Val Ala Asn Tyr Ser
    210                 215                 220
Leu Asp Gly Leu Arg Ile Asp Ala Ala Lys His Val Asp Lys Pro Phe
225                 230                 235                 240
Trp Pro Gly Phe Gln Gln Ala Ala Asn Thr Phe Thr Ile Gly Glu Val
                245                 250                 255
Phe Asp Glu Leu Pro Thr Asn Ala Cys Glu Trp Ala Val Asp Ala Leu
            260                 265                 270
Ser Ser Val Leu Asn Tyr Pro Ala Trp Tyr Tyr Ile Thr Ser Ile Leu
        275                 280                 285
Ser Asn Ser Thr Asn Gly Met Gly Ser Leu Ser Tyr Gln Phe Asp Gln
    290                 295                 300
Thr Gln Gln Tyr Cys Tyr Asp Thr Thr Leu Leu Gly Thr Phe Ser Glu
305                 310                 315                 320
Asn His Asp Val Ala Arg Phe Gly Ser Tyr Thr Ser Asp Val Ser Gln
                325                 330                 335
Arg Lys Asn Ala Leu Thr Phe Asp Phe Phe Thr Asp Gly Ile Pro Ile
            340                 345                 350
Val Tyr Tyr Gly Ala Glu Gln Gly Leu Asp Gly Glu Asn Asp Pro Gln
        355                 360                 365
Asn Arg Gly Ala Leu Trp Leu Asn Pro Thr Gly Tyr Asp Thr Thr Ala
    370                 375                 380
Val Leu Tyr Gln His Ile Lys Thr Met Asn Thr Ala Arg Asn Ala Val
385                 390                 395                 400
Asn Asn Tyr Met Ile Ala Thr Asn Tyr Ser Asn Trp Ser Pro Tyr Trp
                405                 410                 415
Ala Tyr Lys Ala Ser Val Ile Lys Gln Ala Asp Asp Val Leu Val Phe
            420                 425                 430
Arg Lys Gly Leu Val His Ser Ile Val Thr Ala Ile Thr Asn Val Gly
        435                 440                 445
Thr Glu Gly Ala Thr Val Gly Pro Tyr Tyr Ile Glu Asp Thr Asn Phe
    450                 455                 460
Ser Glu Gly Asn Leu Leu Ile Glu Ile Leu Ser Cys Asn Ser Thr Val
465                 470                 475                 480
Ala Gly Val Gly Gly Ala Phe Thr Leu Thr Leu Thr Ser Gly Glu Pro
                485                 490                 495
Gln Val Cys Leu Tyr Ser Phe Asn Ala Tyr Ala
            500                 505

<210> 567
<211> 1446
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 567
atggatttaa cggcgccgaa gcaagccagg ggcaattgca tttgccaccc catggtttgt      60
ttttcggtga gcttcaacag gagtaagcca gtgcttcgtc gatacttgtt cctggtattt     120
attctgctgc taagcggcgc caccaaagca gaagcggctt ttgtacattt gtttgaatgg     180
cgttggctgg atatcgccag cgagtgtgaa aattttctcg gccccaaggg ttttagcgcg     240
gtgcagattt caccaccgca agagcatctg gatcatcaca cctggtgggc ccgctaccag     300
ccggtatcgc tgacaaagct aaccagccgt agcggcacag aaacagaatt gcgcgagatg     360
gttgcgcgct gtgctgccgc aggtgtaaaa atttatgccg atgtggtcat caacaactgg     420
gctactttac ccggcgccag ccgcataggt agtgcgggta gccagtggga agcctaccaa     480
tacccggatc tggggccgga tgattttcat tcgccgcgct gtggtataga aacctatcaa     540
aacgcacatc aggtgtggta ctgcgaactc tacggcatgc cggatttata cactggcgcg     600
cccggccccc aagcctatgt ggcagattac atcaagcgtt taacggcgat gggcatcgcc     660
ggttttcgta tggatgcggc caaacatatt gcgccagcgg atttgcaaca gattcttgcc     720
caggccgaca acccctgggt ttttgcggag gttataggca cgccgggtga agcgccggaa     780
attcacccgc gccaatatct atcactcggt catgttactg agtttgcctt tggcaatctg     840
ctcgcccagg cgtttaacgg ccgcgtggat ctattgcgtg atctgcccaa aagtgattgg     900
ctgccgagcg aacaagcagt gatttttgtc gataaccatg atcgcgaacg cggccacggt     960
ggtgagggca atctcactta caaggatggc agccgctaca agttggccaa tgtgtttatg    1020
```

ctggcccagc cctacggtta tccggcaatt atgtcgggct atgaatttac cgatacagat 1080
gccggccctc cgcccggtga aactcattgc gacaatcgcg cctgggtatg ccagcaccgc 1140
tggccagaag tggcgaatat gcttgggttt cgcgctcaca cgcaaaatga acctttgctg 1200
cactggtggg acaatggccg cgatcaacta gcctttgcgc gcggcaagcg cggatttgtg 1260
gcaattaaca acagcagcaa gcgactcaat cgacgcatgc aaaccggatt gcctgctggt 1320
cgttattgca atcaattggc agagaagtcc gcctgcaacg cgccaatcaa cgtagacaaa 1380
aaaggtcgtg cctggattcg tgtaaagcca ggtaatgcat tggtgttgat taatctgtca 1440
acctag 1446

<210> 568
<211> 481
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> DOMAIN
<222> (52)...(391)
<223> Alpha amylase, catalytic domain

<221> DOMAIN
<222> (400)...(481)
<223> Alpha amylase, C-terminal all-beta domain

<400> 568
Met Asp Leu Thr Ala Pro Lys Gln Ala Arg Gly Asn Cys Ile Cys His
 1   5   10   15
Pro Met Val Cys Phe Ser Val Ser Phe Asn Arg Ser Lys Pro Val Leu
            20   25   30
Arg Arg Tyr Leu Phe Leu Val Phe Ile Leu Leu Leu Ser Gly Ala Thr
        35   40   45
Lys Ala Glu Ala Ala Phe Val His Leu Phe Glu Trp Arg Trp Leu Asp
    50   55   60
Ile Ala Ser Glu Cys Glu Asn Phe Leu Gly Pro Lys Gly Phe Ser Ala
65   70   75   80
Val Gln Ile Ser Pro Pro Gln Glu His Leu Asp His His Thr Trp Trp
                85   90   95
Ala Arg Tyr Gln Pro Val Ser Leu Thr Lys Leu Thr Ser Arg Ser Gly
            100   105   110
Thr Glu Thr Glu Leu Arg Glu Met Val Ala Arg Cys Ala Ala Ala Gly
        115   120   125
Val Lys Ile Tyr Ala Asp Val Val Ile Asn Asn Trp Ala Thr Leu Pro
    130   135   140
Gly Ala Ser Arg Ile Gly Ser Ala Gly Ser Gln Trp Glu Ala Tyr Gln
145   150   155   160
Tyr Pro Asp Leu Gly Pro Asp Asp Phe His Ser Pro Arg Cys Gly Ile
                165   170   175
Glu Thr Tyr Gln Asn Ala His Gln Val Trp Tyr Cys Glu Leu Tyr Gly
            180   185   190
Met Pro Asp Leu Tyr Thr Gly Ala Pro Gly Pro Gln Ala Tyr Val Ala
        195   200   205
Asp Tyr Ile Lys Arg Leu Thr Ala Met Gly Ile Ala Gly Phe Arg Met
    210   215   220
Asp Ala Ala Lys His Ile Ala Pro Ala Asp Leu Gln Gln Ile Leu Ala
225   230   235   240
Gln Ala Asp Asn Pro Trp Val Phe Ala Glu Val Ile Gly Thr Pro Gly
                245   250   255
Glu Ala Pro Glu Ile His Pro Arg Gln Tyr Leu Ser Leu Gly His Val
            260   265   270
Thr Glu Phe Ala Phe Gly Asn Leu Leu Ala Gln Ala Phe Asn Gly Arg
        275   280   285
Val Asp Leu Leu Arg Asp Leu Pro Lys Ser Asp Trp Leu Pro Ser Glu
    290   295   300
Gln Ala Val Ile Phe Val Asp Asn His Asp Arg Glu Arg Gly His Gly
305   310   315   320
Gly Glu Gly Asn Leu Thr Tyr Lys Asp Gly Ser Arg Tyr Lys Leu Ala
                325   330   335

```
Asn Val Phe Met Leu Ala Gln Pro Tyr Gly Tyr Pro Ala Ile Met Ser
            340                 345                 350
Gly Tyr Glu Phe Thr Asp Thr Asp Ala Gly Pro Pro Pro Gly Glu Thr
        355                 360                 365
His Cys Asp Asn Arg Ala Trp Val Cys Gln His Arg Trp Pro Glu Val
    370                 375                 380
Ala Asn Met Leu Gly Phe Arg Ala His Thr Gln Asn Glu Pro Leu Leu
385                 390                 395                 400
His Trp Trp Asp Asn Gly Arg Asp Gln Leu Ala Phe Ala Arg Gly Lys
                405                 410                 415
Arg Gly Phe Val Ala Ile Asn Asn Ser Ser Lys Arg Leu Asn Arg Arg
            420                 425                 430
Met Gln Thr Gly Leu Pro Ala Gly Arg Tyr Cys Asn Gln Leu Ala Glu
        435                 440                 445
Lys Ser Ala Cys Asn Ala Pro Ile Asn Val Asp Lys Lys Gly Arg Ala
    450                 455                 460
Trp Ile Arg Val Lys Pro Gly Asn Ala Leu Val Leu Ile Asn Leu Ser
465                 470                 475                 480
Thr
```

<210> 569
<211> 1095
<212> DNA
<213> Thermococcus alcaliphilus AEDII12RA

<400> 569

```
ttgagagcgc tcgtctttca cggcaacctc cagtatgccg aaatcccaaa gagcgaaatc      60
ccaaaggtca tagagaaggc atacatccca gtcatcgaga cactgattaa agaagaaatt     120
ccttttgggc tcaacataac gggctatacc ttaaagttcc tcccgaagga tattatagac     180
ctcgttaaag ggggcatcgc gagtgacctg atagagataa tcggaacgag ctacacgcac     240
gcaatactcc ccctcctgcc gcttagcaga gtagaagcac aagttcagag agatagggaa     300
gttaaggaag agctcttcga ggtttctcca aagggattct ggctgccaga gctcgcctat     360
gacccgataa tccctgccat actgaaggac aacggttatg agtatctatt cgccgacggg     420
gaggcgatgc ttttctcagc tcatctcaac tcggcgataa agccaattaa accgctctat     480
ccacacctta taaaggccca aagggaaaag cgctttaggt acatcagcta tctccttggt     540
ctcagggagc ttaggaaggc gataaagctc gtttttgaag gtaaggtaac gctaaaggca     600
gtcaaagaca tcgaagccgt acccgtttgg gtggccgtga acacggctgt aatgctcggc     660
atcggaaggc ttcctcttat gaatcctaag aaagtggcga gctggataga ggacaaggac     720
aacattcttc tatacggcac cgatatagag ttcattggct atagggacat tgcaggctac     780
agaatgagtg ttgagggatt attagaggtt atagacgagc tcaactcgga actgtgcctt     840
ccctcagagc tgaagcacag tggaagggag ctctacttac ggacttcgag ttgggcacca     900
gataagagct tgaggatatg gagagaggac gaagggaacg caagacttaa tatgctgtcc     960
tacaatatga ggggcgaact cgcctttta gccgagaaca gcgatgcaag gggatgggag    1020
cccctccctg agaggaggct ggatgccttc cgggcgatat ataacgattg gaggggtgaa    1080
aatggggaac cttag                                                    1095
```

<210> 570
<211> 364
<212> PRT
<213> Thermococcus alcaliphilus AEDII12RA

<400> 570

```
Met Arg Ala Leu Val Phe His Gly Asn Leu Gln Tyr Ala Glu Ile Pro
 1               5                  10                  15
Lys Ser Glu Ile Pro Lys Val Ile Glu Lys Ala Tyr Ile Pro Val Ile
            20                  25                  30
Glu Thr Leu Ile Lys Glu Glu Ile Pro Phe Gly Leu Asn Ile Thr Gly
        35                  40                  45
Tyr Thr Leu Lys Phe Leu Pro Lys Asp Ile Ile Asp Leu Val Lys Gly
    50                  55                  60
Gly Ile Ala Ser Asp Leu Ile Glu Ile Ile Gly Thr Ser Tyr Thr His
65                  70                  75                  80
Ala Ile Leu Pro Leu Leu Pro Leu Ser Arg Val Glu Ala Gln Val Gln
                85                  90                  95
Arg Asp Arg Glu Val Lys Glu Glu Leu Phe Glu Val Ser Pro Lys Gly
            100                 105                 110
Phe Trp Leu Pro Glu Leu Ala Tyr Asp Pro Ile Ile Pro Ala Ile Leu
```

```
        115                 120                 125
Lys Asp Asn Gly Tyr Glu Tyr Leu Phe Ala Asp Gly Glu Ala Met Leu
    130                 135                 140
Phe Ser Ala His Leu Asn Ser Ala Ile Lys Pro Ile Lys Pro Leu Tyr
145                 150                 155                 160
Pro His Leu Ile Lys Ala Gln Arg Glu Lys Arg Phe Arg Tyr Ile Ser
                165                 170                 175
Tyr Leu Leu Gly Leu Arg Glu Leu Arg Lys Ala Ile Lys Leu Val Phe
            180                 185                 190
Glu Gly Lys Val Thr Leu Lys Ala Val Lys Asp Ile Glu Ala Val Pro
        195                 200                 205
Val Trp Val Ala Val Asn Thr Ala Val Met Leu Gly Ile Gly Arg Leu
    210                 215                 220
Pro Leu Met Asn Pro Lys Lys Val Ala Ser Trp Ile Glu Asp Lys Asp
225                 230                 235                 240
Asn Ile Leu Leu Tyr Gly Thr Asp Ile Glu Phe Ile Gly Tyr Arg Asp
                245                 250                 255
Ile Ala Gly Tyr Arg Met Ser Val Glu Gly Leu Leu Glu Val Ile Asp
            260                 265                 270
Glu Leu Asn Ser Glu Leu Cys Leu Pro Ser Glu Leu Lys His Ser Gly
        275                 280                 285
Arg Glu Leu Tyr Leu Arg Thr Ser Ser Trp Ala Pro Asp Lys Ser Leu
    290                 295                 300
Arg Ile Trp Arg Glu Asp Glu Gly Asn Ala Arg Leu Asn Met Leu Ser
305                 310                 315                 320
Tyr Asn Met Arg Gly Glu Leu Ala Phe Leu Ala Glu Asn Ser Asp Ala
                325                 330                 335
Arg Gly Trp Glu Pro Leu Pro Glu Arg Arg Leu Asp Ala Phe Arg Ala
            340                 345                 350
Ile Tyr Asn Asp Trp Arg Gly Glu Asn Gly Glu Pro
        355                 360
```

<210> 571
<211> 1839
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 571

```
atgcaatcga acggaaacgt gaaggggagg agcgccgttc tagcgctggc cctcctcctg      60
ttgacagcgg ttgccgccac ggcccagacc aacggcacga tgatgcagta cttccactgg     120
tacaacagct cgtcagacaa cctctggctc aaggtgaccg accaggccga cgagctgtcc     180
gcggccggca tcaccgccct ctggctgccc cccgcctaca aggcgatcaa cggcgacgac     240
gtggggtatg gcgtctacga tctctacgac ctcggcgagt tcaaccagaa ggggtcggtg     300
cgcaccaaat acggcaccag ggcccagtac ctcgccgcca tcaacgcggc tcacgccaac     360
ggcctgcagg tctacggcga cgtggtcttc aaccacaagg gcggcgcgga cttcaccgag     420
tgggtgacgg cgacccgggt gcagacctgg gaccgcaacc aggagtacgg aggcgatctc     480
tccatccagg cctggaccgg cttcgacttc cccggccggg gcaacaccta ctccgcgttc     540
aaatggcgct ggtaccactt cgacggcgtg gactgggacc agaacctcca ggagagcggg     600
aagatctaca agttccgcgg caccggcaag gcctgggact ggaaggtcca cgacgagaac     660
ggcaactacg actatctgat gttcgccgac ctcgacatgc agcacccgga cgtcgtcaac     720
gagctcaaga gctggggcgt ctggtacgcc aacaccaccg gggtcgacgg cttccggctg     780
gatgccctca agcacatcaa gtacgacttc tggaacggct ggctggacca cgtgcgcaac     840
accacgggca agccgctgtt caccgtcggc gagctctgga gctacgacgt ccagaacctg     900
cacgactgga tcacggagac cggcggccgg gcctcgctgt tcgacgcgcc gctgcacctg     960
aacttcttca acgcgtcgcg ggccggcggg agctatgaca tgcgcgacct gatggacggc    1020
acgctgatgg agcagcagcc ggccctggca gtgacgctgg tcgagaacca cgacacgcag    1080
cctctccagg ccctggagtc gccggtgatg gactggttca agcccctggc ctacgccttc    1140
atcctgctgc gcgcggaagg ctacccttcc gtcttctacc cggactacta cggcgcgacc    1200
tatacggaca gaggccagac cgtcaccatc gcctcccaga aggcgatcct cgacaagctg    1260
ctggacgccc gcaggaactt cgcctacggg acgcagtacg actacctcga tcactgggac    1320
gtcatcggct ggacgcgcct gggcaacgcc cagcatcccc gggccatggc cgtgatcctg    1380
agcgacggcc ccggtggctc gaagtggatg gacgtgggca aggccaacgc ccgattcttc    1440
gatcacaccg gcaaccgcgt ggacgtggtg accaccaacg ggtcgggctg ggggaacttc    1500
ccggtcaacg gcggctcggt ctccgtctgg gtgcaggagg accccggcca ggtgtcggtc    1560
ctgttcacct gcaacaacgg ccagacggtg atgggccagg acgtatacgt gacgggcagc    1620
```

gtcgccaaat tgggcaactg ggacacgagc aaggccgtga agctctcgcc ggtgagctac 1680
ccgacctgga gcgggtgggt gcacaacctc cctccgagca cccagatcca atggaagtgc 1740
atcaagaagc aggggacgaa cgtcgtctgg cagccgggcg ccaacaacgt cgtcacgacc 1800
ccggccagcg gctcgacgag ctcgaccggg gcgttctga 1839

<210> 572
<211> 612
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<221> SIGNAL
<222> (1)...(28)

<221> DOMAIN
<222> (517)...(610)
<223> Starch binding domain

<221> DOMAIN
<222> (32)...(424)
<223> Alpha amylase, catalytic domain

<400> 572
Met Gln Ser Asn Gly Asn Val Lys Gly Arg Ser Ala Val Leu Ala Leu
 1   5   10   15
Ala Leu Leu Leu Leu Thr Ala Val Ala Ala Thr Ala Gln Thr Asn Gly
            20   25   30
Thr Met Met Gln Tyr Phe His Trp Tyr Asn Ser Ser Ser Asp Asn Leu
        35   40   45
Trp Leu Lys Val Thr Asp Gln Ala Asp Glu Leu Ser Ala Ala Gly Ile
    50   55   60
Thr Ala Leu Trp Leu Pro Pro Ala Tyr Lys Ala Ile Asn Gly Asp Asp
65   70   75   80
Val Gly Tyr Gly Val Tyr Asp Leu Tyr Asp Leu Gly Glu Phe Asn Gln
                85   90   95
Lys Gly Ser Val Arg Thr Lys Tyr Gly Thr Arg Ala Gln Tyr Leu Ala
            100   105   110
Ala Ile Asn Ala Ala His Ala Asn Gly Leu Gln Val Tyr Gly Asp Val
        115   120   125
Val Phe Asn His Lys Gly Gly Ala Asp Phe Thr Glu Trp Val Thr Ala
    130   135   140
Thr Arg Val Gln Thr Trp Asp Arg Asn Gln Glu Tyr Gly Gly Asp Leu
145   150   155   160
Ser Ile Gln Ala Trp Thr Gly Phe Asp Phe Pro Gly Arg Gly Asn Thr
                165   170   175
Tyr Ser Ala Phe Lys Trp Arg Trp Tyr His Phe Asp Gly Val Asp Trp
            180   185   190
Asp Gln Asn Leu Gln Glu Ser Gly Lys Ile Tyr Lys Phe Arg Gly Thr
        195   200   205
Gly Lys Ala Trp Asp Trp Lys Val His Asp Glu Asn Gly Asn Tyr Asp
    210   215   220
Tyr Leu Met Phe Ala Asp Leu Asp Met Gln His Pro Asp Val Val Asn
225   230   235   240
Glu Leu Lys Ser Trp Gly Val Trp Tyr Ala Asn Thr Thr Gly Val Asp
                245   250   255
Gly Phe Arg Leu Asp Ala Leu Lys His Ile Lys Tyr Asp Phe Trp Asn
            260   265   270
Gly Trp Leu Asp His Val Arg Asn Thr Thr Gly Lys Pro Leu Phe Thr
        275   280   285
Val Gly Glu Leu Trp Ser Tyr Asp Val Gln Asn Leu His Asp Trp Ile
    290   295   300
Thr Glu Thr Gly Gly Arg Ala Ser Leu Phe Asp Ala Pro Leu His Leu
305   310   315   320
Asn Phe Phe Asn Ala Ser Arg Ala Gly Gly Ser Tyr Asp Met Arg Asp
                325   330   335
Leu Met Asp Gly Thr Leu Met Glu Gln Gln Pro Ala Leu Ala Val Thr

```
            340                 345                 350
Leu Val Glu Asn His Asp Thr Gln Pro Leu Gln Ala Leu Glu Ser Pro
        355                 360                 365
Val Met Asp Trp Phe Lys Pro Leu Ala Tyr Ala Phe Ile Leu Leu Arg
    370                 375                 380
Ala Glu Gly Tyr Pro Ser Val Phe Tyr Pro Asp Tyr Tyr Gly Ala Thr
385                 390                 395                 400
Tyr Thr Asp Arg Gly Gln Thr Val Thr Ile Ala Ser Gln Lys Ala Ile
                405                 410                 415
Leu Asp Lys Leu Leu Asp Ala Arg Arg Asn Phe Ala Tyr Gly Thr Gln
            420                 425                 430
Tyr Asp Tyr Leu Asp His Trp Asp Val Ile Gly Trp Thr Arg Leu Gly
        435                 440                 445
Asn Ala Gln His Pro Arg Ala Met Ala Val Ile Leu Ser Asp Gly Pro
    450                 455                 460
Gly Gly Ser Lys Trp Met Asp Val Gly Lys Ala Asn Ala Arg Phe Phe
465                 470                 475                 480
Asp His Thr Gly Asn Arg Val Asp Val Val Thr Thr Asn Gly Ser Gly
                485                 490                 495
Trp Gly Asn Phe Pro Val Asn Gly Gly Ser Val Ser Val Trp Val Gln
            500                 505                 510
Glu Asp Pro Gly Gln Val Ser Val Leu Phe Thr Cys Asn Asn Gly Gln
        515                 520                 525
Thr Val Met Gly Gln Asp Val Tyr Val Thr Gly Ser Val Ala Lys Leu
    530                 535                 540
Gly Asn Trp Asp Thr Ser Lys Ala Val Lys Leu Ser Pro Val Ser Tyr
545                 550                 555                 560
Pro Thr Trp Ser Gly Trp Val His Asn Leu Pro Pro Ser Thr Gln Ile
                565                 570                 575
Gln Trp Lys Cys Ile Lys Lys Gln Gly Thr Asn Val Val Trp Gln Pro
            580                 585                 590
Gly Ala Asn Asn Val Val Thr Thr Pro Ala Ser Gly Ser Thr Ser Ser
        595                 600                 605
Thr Gly Ala Phe
    610
```

<210> 573
<211> 2133
<212> DNA
<213> Bacteria

<400> 573

```
atgaaaaaaa cgtttaaatt gatattggtg ctgatgcttt cacttacact tgtttttgga      60
ttgacagcac caatacaggc agcttctgat acagcagtaa gtaatgttgt caattactcg     120
acggatgtaa tttatcagat tgtcactgat agatttgttg atggaaatac cagcaacaat     180
ccaactggtg atttgtatga tccaacacat acaagtctga aaaagtactt tggtggagac     240
tggcagggta ttattaacaa gataaatgat ggttatctaa caggaatggg tgtaacagca     300
atttggatat ctcagcctgt agaaaatatt tatgcagtct tacctgattc aacttttggt     360
ggaagtacat cgtaccatgg ttattgggct cgtgatttta agagaacaaa tccatacttt     420
ggaagcttta ctgattttca aaatttgatt aatacagctc atgcacataa cataaaagta     480
attatcgatt ttgcaccaaa tcatacttca cctgcatctg aaaccgatcc gacttatgct     540
gaaaatggaa ggctatacga caatggaaca ttacttggtg ggtatacaaa tgatacaaac     600
gggtattttc atcattatgg aggaacagat ttttcatctt atgaagatgg catttatcgt     660
aacttatttg atttagcaga tttaaatcaa cagaatagta caatcgattc atacttaaaa     720
tcagcaatta aggtgtggct cgatatggga atagacggta tacgtctaga tgctgtaaaa     780
catatgccgt ttggatggca gaagaatttt atggatagca tattgagcta taggcctgta     840
tttacatttg gagagtggtt tcttggaacg aatgaaattg atgtaaataa cacatacttt     900
gcaaatgaaa gcggtatgag tcttttggac tttaggtttt ctcaaaaagt gaggcaggta     960
tttagggata atacggatac tatgtatgga cttgattcaa tgattcaatc aactgcatct    1020
gattataatt ttataaatga tatggttact tttattgata atcatgatat ggatagattc    1080
tataatggtg gttctactcg tccagttgaa caagcattag catttacttt gacttcgagg    1140
ggagtacctg ctatatatta tggaacagaa cagtatatga caggcaatgg agacccttat    1200
aatagagcta tgatgacctc atttaataca tcaactacag catataatgt aattaaaaaa    1260
cttgctcctt tgcgcaaatc taatcctgca attgcttatg gtactactca gcaacgttgg    1320
ataaacaatg atgtttatat atatgaaaga aaatttggga ataatgttgc tcttgttgct    1380
ataaatagaa atctttcaac aagttataat ataacaggac tgtatacagc gcttcctgct    1440
ggtacttata ctgatgttct tggtggactt ttaaatggta atagtatttc tgtcgcgagt    1500
gatggttctg taacaccatt tacacttagt gccggtgaag ttgcagtatg gcagtatgta    1560
```

agttcaagta attctccgtt gataggacat gttgggccaa caatgacaaa agcagggcaa 1620
actataacaa tagatggaag aggttttggt acgacatcag gacaagtttt gtttggttcg 1680
acggctggta ctattgtctc ttgggatgat acagaggtaa aagtaaaagt tccttcggtg 1740
acacctggta aatataatat atcattaaaa acttcaagtg gtgctacaag caatacttat 1800
aataatatta acattcttac aggcaatcag atatgtgtta gatttgttgt aaataatgca 1860
agtacagtgt atggagagaa tgtttatctt actggtaatg tagctgaatt gggaaactgg 1920
gatacgtcaa aagcaatagg accaatgttt aatcaggtag tatatcagta tcctacatgg 1980
tattatgatg taagtgttcc agcgggtaca acaatacaat ttaagtttat aaagaaaaat 2040
ggcaatacga taacatggga aggtggtagt aaccatactt atacagttcc atcatctagt 2100
acaggtactg taattgtaaa ctggcaacaa tag 2133

<210> 574
<211> 710
<212> PRT
<213> Bacteria

<220>
<221> SIGNAL
<222> (1)...(27)

<221> DOMAIN
<222> (437)...(521)
<223> Alpha amylase, C-terminal all-beta domain

<221> DOMAIN
<222> (46)...(425)
<223> Alpha amylase, catalytic domain

<221> DOMAIN
<222> (526)...(603)
<223> IPT/TIG domain

<221> DOMAIN
<222> (610)...(706)
<223> Starch binding domain

<400> 574
Met Lys Lys Thr Phe Lys Leu Ile Leu Val Leu Met Leu Ser Leu Thr
 1   5   10   15
Leu Val Phe Gly Leu Thr Ala Pro Ile Gln Ala Ala Ser Asp Thr Ala
            20   25   30
Val Ser Asn Val Val Asn Tyr Ser Thr Asp Val Ile Tyr Gln Ile Val
        35   40   45
Thr Asp Arg Phe Val Asp Gly Asn Thr Ser Asn Asn Pro Thr Gly Asp
    50   55   60
Leu Tyr Asp Pro Thr His Thr Ser Leu Lys Lys Tyr Phe Gly Gly Asp
65   70   75   80
Trp Gln Gly Ile Ile Asn Lys Ile Asn Asp Gly Tyr Leu Thr Gly Met
                85   90   95
Gly Val Thr Ala Ile Trp Ile Ser Gln Pro Val Glu Asn Ile Tyr Ala
            100   105   110
Val Leu Pro Asp Ser Thr Phe Gly Gly Ser Thr Ser Tyr His Gly Tyr
        115   120   125
Trp Ala Arg Asp Phe Lys Arg Thr Asn Pro Tyr Phe Gly Ser Phe Thr
    130   135   140
Asp Phe Gln Asn Leu Ile Asn Thr Ala His Ala His Asn Ile Lys Val
145   150   155   160
Ile Ile Asp Phe Ala Pro Asn His Thr Ser Pro Ala Ser Glu Thr Asp
                165   170   175
Pro Thr Tyr Ala Glu Asn Gly Arg Leu Tyr Asp Asn Gly Thr Leu Leu
            180   185   190
Gly Gly Tyr Thr Asn Asp Thr Asn Gly Tyr Phe His His Tyr Gly Gly
        195   200   205
Thr Asp Phe Ser Ser Tyr Glu Asp Gly Ile Tyr Arg Asn Leu Phe Asp
    210   215   220
Leu Ala Asp Leu Asn Gln Gln Asn Ser Thr Ile Asp Ser Tyr Leu Lys
225   230   235   240
Ser Ala Ile Lys Val Trp Leu Asp Met Gly Ile Asp Gly Ile Arg Leu

```
                245                 250                 255
Asp Ala Val Lys His Met Pro Phe Gly Trp Gln Lys Asn Phe Met Asp
            260                 265                 270
Ser Ile Leu Ser Tyr Arg Pro Val Phe Thr Phe Gly Glu Trp Phe Leu
        275                 280                 285
Gly Thr Asn Glu Ile Asp Val Asn Asn Thr Tyr Phe Ala Asn Glu Ser
    290                 295                 300
Gly Met Ser Leu Leu Asp Phe Arg Phe Ser Gln Lys Val Arg Gln Val
305                 310                 315                 320
Phe Arg Asp Asn Thr Asp Thr Met Tyr Gly Leu Asp Ser Met Ile Gln
                325                 330                 335
Ser Thr Ala Ser Asp Tyr Asn Phe Ile Asn Asp Met Val Thr Phe Ile
            340                 345                 350
Asp Asn His Asp Met Asp Arg Phe Tyr Asn Gly Gly Ser Thr Arg Pro
        355                 360                 365
Val Glu Gln Ala Leu Ala Phe Thr Leu Thr Ser Arg Gly Val Pro Ala
    370                 375                 380
Ile Tyr Tyr Gly Thr Glu Gln Tyr Met Thr Gly Asn Gly Asp Pro Tyr
385                 390                 395                 400
Asn Arg Ala Met Met Thr Ser Phe Asn Thr Ser Thr Thr Ala Tyr Asn
                405                 410                 415
Val Ile Lys Lys Leu Ala Pro Leu Arg Lys Ser Asn Pro Ala Ile Ala
            420                 425                 430
Tyr Gly Thr Thr Gln Gln Arg Trp Ile Asn Asn Asp Val Tyr Ile Tyr
        435                 440                 445
Glu Arg Lys Phe Gly Asn Asn Val Ala Leu Val Ala Ile Asn Arg Asn
    450                 455                 460
Leu Ser Thr Ser Tyr Asn Ile Thr Gly Leu Tyr Thr Ala Leu Pro Ala
465                 470                 475                 480
Gly Thr Tyr Thr Asp Val Leu Gly Gly Leu Leu Asn Gly Asn Ser Ile
                485                 490                 495
Ser Val Ala Ser Asp Gly Ser Val Thr Pro Phe Thr Leu Ser Ala Gly
            500                 505                 510
Glu Val Ala Val Trp Gln Tyr Val Ser Ser Ser Asn Ser Pro Leu Ile
        515                 520                 525
Gly His Val Gly Pro Thr Met Thr Lys Ala Gly Gln Thr Ile Thr Ile
    530                 535                 540
Asp Gly Arg Gly Phe Gly Thr Thr Ser Gly Gln Val Leu Phe Gly Ser
545                 550                 555                 560
Thr Ala Gly Thr Ile Val Ser Trp Asp Asp Thr Glu Val Lys Val Lys
                565                 570                 575
Val Pro Ser Val Thr Pro Gly Lys Tyr Asn Ile Ser Leu Lys Thr Ser
            580                 585                 590
Ser Gly Ala Thr Ser Asn Thr Tyr Asn Asn Ile Asn Ile Leu Thr Gly
        595                 600                 605
Asn Gln Ile Cys Val Arg Phe Val Val Asn Asn Ala Ser Thr Val Tyr
    610                 615                 620
Gly Glu Asn Val Tyr Leu Thr Gly Asn Val Ala Glu Leu Gly Asn Trp
625                 630                 635                 640
Asp Thr Ser Lys Ala Ile Gly Pro Met Phe Asn Gln Val Val Tyr Gln
                645                 650                 655
Tyr Pro Thr Trp Tyr Tyr Asp Val Ser Val Pro Ala Gly Thr Thr Ile
            660                 665                 670
Gln Phe Lys Phe Ile Lys Lys Asn Gly Asn Thr Ile Thr Trp Glu Gly
        675                 680                 685
Gly Ser Asn His Thr Tyr Thr Val Pro Ser Ser Ser Thr Gly Thr Val
    690                 695                 700
Ile Val Asn Trp Gln Gln
705                 710

<210> 575
<211> 1677
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.
```

<400> 575
atgaaaagcg agttaccaca aattccacca cagcaggtga acaacaccat gtaccaggca 60
ttttattggg atgcctaccc tggcctttgg gccaatttac cggccatggc ggccccttg 120
gccgaacggg gcattacctc gatgtggttg ccgcccgccg ccaaaggcat gaatggtact 180
ttcagtgtcg gttacgatgt atacgatttc tgggatctgg gcgagtttaa ccaaaaaggc 240
accaccgcca cccgttacgg tactcgtcag cagctgcaac aagcactgag tgctctggac 300
caactgggta ttcaggccta ttttgatgtg gtgtttaacc accgcatggg cgccgatgca 360
caggagaata ttcctggctt tggcctggcc tggaccgagt atcatctgca aggtcgtcag 420
gcgcattata cccagcaaaa ctggggctac ttgtggcacg actttgactg gaactggacc 480
gcgtttaatg gctccgacaa tcagctctac cccggcaaat ggtggggcaa taccttccac 540
ttcccttatt tgatgggtga ggatgtcgat tacaaccgct ttgaagtgca gcaggaaatg 600
aaagcctggg gcgagtggat catcaacagc gttggcttta gcggctttcg gatggacgcc 660
atcgcacacg tggataccga ttttacccgt gactggatca atcacgtaca gtgggccacc 720
agcgaggatg tattctttgt cgcagaagcc tgggtcagtg atatcaacgg ctatctggat 780
gcagtcaata cgccgcattt gcgcgctttt gatttcaatt tgcgcgaaga ctttgttgct 840
ttaagcagtg gtggcaaaga catgcgctgg tggggtggtc tggttaatag ccagcaccgt 900
gatcgggcgg tcacttttgt cgataaccac gataccagcc gggccggcaa cccttatggc 960
atgccgcagg tgatcaacta caagaaccag gcctacgctt acattctgtt gcgtgagcat 1020
ggggtgccga ctgtgtttgc ccgcgattac gacgaatttg gcatggcgcc aacgctggat 1080
aaattgattg aggcgcgccg ctactttgcc tatggtcctg gtcatgagta ctccggcaat 1140
accgaggccg tttacgccta tgtgcgcgag gggcttagca ctgtgccggg caccggtctg 1200
gtgatgctga tatcgggtcg aaactggggt ggtcagcagt cgttcaccat caacagccat 1260
cagccgaata ccacctttta cgattatacc ggcaatgtca gcggcacggt gaccaccaat 1320
gcgcagggtt atggcagctt cccggtcact atgacggaaa gtaccggttg gtcagtctgg 1380
gtaccacaat ccagtggtgg cactcagccg ggatccatta ccctgcggat gaccaaggat 1440
gttggctatg gtttttcact gttcttcact ggcagcagtg cagagctcac caattggggc 1500
ggcggtattg aaggcacctg gaccagtggc catgtctggg aagtgaccat cccggatccg 1560
ggcaactttg aatggaaaac ccgtaaaggc ccaaccggcg gcagtggtca ggactgggaa 1620
agtggtagca accacaatca gaccaatttg caccccagtt ttaatggtgg gttttaa 1677

<210> 576
<211> 558
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<221> DOMAIN
<222> (15)...(366)
<223> Alpha amylase, catalytic domain

<400> 576
Met Lys Ser Glu Leu Pro Gln Ile Pro Pro Gln Gln Val Asn Asn Thr
 1   5   10   15
Met Tyr Gln Ala Phe Tyr Trp Asp Ala Tyr Pro Gly Leu Trp Ala Asn
            20   25   30
Leu Pro Ala Met Ala Ala Pro Leu Ala Glu Arg Gly Ile Thr Ser Met
        35   40   45
Trp Leu Pro Pro Ala Ala Lys Gly Met Asn Gly Thr Phe Ser Val Gly
    50   55   60
Tyr Asp Val Tyr Asp Phe Trp Asp Leu Gly Glu Phe Asn Gln Lys Gly
65   70   75   80
Thr Thr Ala Thr Arg Tyr Gly Thr Arg Gln Gln Leu Gln Gln Ala Leu
                85   90   95
Ser Ala Leu Asp Gln Leu Gly Ile Gln Ala Tyr Phe Asp Val Val Phe
            100   105   110
Asn His Arg Met Gly Ala Asp Ala Gln Glu Asn Ile Pro Gly Phe Gly
        115   120   125
Leu Ala Trp Thr Glu Tyr His Leu Gln Gly Arg Gln Ala His Tyr Thr
    130   135   140
Gln Gln Asn Trp Gly Tyr Leu Trp His Asp Phe Asp Trp Asn Trp Thr
145   150   155   160
Ala Phe Asn Gly Ser Asp Asn Gln Leu Tyr Pro Gly Lys Trp Trp Gly
                165   170   175
Asn Thr Phe His Phe Pro Tyr Leu Met Gly Glu Asp Val Asp Tyr Asn
            180   185   190
Arg Phe Glu Val Gln Gln Glu Met Lys Ala Trp Gly Glu Trp Ile Ile

```
        195                 200                 205
Asn Ser Val Gly Phe Ser Gly Phe Arg Met Asp Ala Ile Ala His Val
    210                 215                 220
Asp Thr Asp Phe Thr Arg Asp Trp Ile Asn His Val Gln Trp Ala Thr
225                 230                 235                 240
Ser Glu Asp Val Phe Phe Val Ala Glu Ala Trp Val Ser Asp Ile Asn
                245                 250                 255
Gly Tyr Leu Asp Ala Val Asn Thr Pro His Leu Arg Ala Phe Asp Phe
            260                 265                 270
Asn Leu Arg Glu Asp Phe Val Ala Leu Ser Ser Gly Gly Lys Asp Met
        275                 280                 285
Arg Trp Trp Gly Gly Leu Val Asn Ser Gln His Arg Asp Arg Ala Val
    290                 295                 300
Thr Phe Val Asp Asn His Asp Thr Ser Arg Ala Gly Asn Pro Tyr Gly
305                 310                 315                 320
Met Pro Gln Val Ile Asn Tyr Lys Asn Gln Ala Tyr Ala Tyr Ile Leu
                325                 330                 335
Leu Arg Glu His Gly Val Pro Thr Val Phe Ala Arg Asp Tyr Asp Glu
            340                 345                 350
Phe Gly Met Ala Pro Thr Leu Asp Lys Leu Ile Glu Ala Arg Arg Tyr
        355                 360                 365
Phe Ala Tyr Gly Pro Gly His Glu Tyr Ser Gly Asn Thr Glu Ala Val
    370                 375                 380
Tyr Ala Tyr Val Arg Glu Gly Leu Ser Thr Val Pro Gly Thr Gly Leu
385                 390                 395                 400
Val Met Leu Ile Ser Gly Arg Asn Trp Gly Gly Gln Gln Ser Phe Thr
                405                 410                 415
Ile Asn Ser His Gln Pro Asn Thr Thr Phe Tyr Asp Tyr Thr Gly Asn
            420                 425                 430
Val Ser Gly Thr Val Thr Thr Asn Ala Gln Gly Tyr Gly Ser Phe Pro
        435                 440                 445
Val Thr Met Thr Glu Ser Thr Gly Trp Ser Val Trp Val Pro Gln Ser
    450                 455                 460
Ser Gly Gly Thr Gln Pro Gly Ser Ile Thr Leu Arg Met Thr Lys Asp
465                 470                 475                 480
Val Gly Tyr Gly Phe Ser Leu Phe Phe Thr Gly Ser Ser Ala Glu Leu
                485                 490                 495
Thr Asn Trp Gly Gly Gly Ile Glu Gly Thr Trp Thr Ser Gly His Val
            500                 505                 510
Trp Glu Val Thr Ile Pro Asp Pro Gly Asn Phe Glu Trp Lys Thr Arg
        515                 520                 525
Lys Gly Pro Thr Gly Gly Ser Gly Gln Asp Trp Glu Ser Gly Ser Asn
    530                 535                 540
His Asn Gln Thr Asn Leu His Pro Ser Phe Asn Gly Gly Phe
545                 550                 555
```

<210> 577
<211> 2160
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 577

```
atgaaaccgt ttttaaaaaa atctattatc acactgcttg cttccacctg tctctttaca      60
gcatggctca tcccctcaat cgccgtacct acggtaagcg ccgctcccga tacgtcggtt     120
cttaacaaac agggcttctc aaccgatgta atctatcaaa tcgtgaccga ccgcttcgcg     180
gatggagatc cttccaacaa tccatccggc gcggcctatt ctcccggctg tacgaatctg     240
aagctttatt gcgggggaga ctggcgggga atcattaaca agattaacga gggttatttt     300
gccggaatgg gtgtaagcgc tttatggatc tctcaaccgg tcgaaaatat ttacagcgtc     360
atcgactact ccggtgtcaa cagcacatcc tatcacgggt attgggctcg cgacttcaaa     420
cggacgaatc cggcgtttgg cagcatgagc gattttcagg agctcatcaa cacggcacat     480
gcccacaata tcaagatcat tattgatttc gcacctaacc acacctcccc tgcttcggag     540
acacaacctt ccttcgccga gaacggtcgg ctgtacgata atggcaacct gatcgctggg     600
tacacggggg acaccaacgg aattttccat cataatcagg gcacgaattt ctcctcgctc     660
gaggatggca tctatcggaa tctttatgat ctggccgaca tcaatcacca caacaacgta     720
acggatactt acttcaagga cgcgattaag ctctggctga atatggggat cgatggcatc     780
```

cgggtggatg cggtgaagca tatgccggag ggttggcaga aaaactgggt cgcttcgatc 840
gcagggtata aaccggtgtt tacctttggg gaatggttcc ttggcgttgg cgaaaatgat 900
ccgaacaaca tccgattcgc caatgaatcc gggatgagtc tgctcgactt ccaatttggg 960
caaaaggtcc ggcaggtctt ccgcgacaat tcagataata tgtacggttt aaataacatg 1020
ctaaccacaa cagcagctaa ctacaaacat attcatgacc aagtcacgtt tatcgacaac 1080
catgatatgg atcgttttaa gctgggtaat aatgaccgtc gcttggagca agcgttggcc 1140
ttcacgctga tgtcacgcgg cgttccggcc atctattacg ggacggagca atatatgacg 1200
ggtaacgggg atccggccaa ccgtggtttt atgagtggat tttcaacttc aaccaccgcc 1260
tatcaggtca tcggcaagct ggcaccgctt cgcaaatcga atccggcact ggcatatgga 1320
agtacgcagg aacgctggat caataatgat gtatttatct tcgagcgtaa gttcggcagc 1380
agcgtagcgc tagtcgcaat taaccgcagc acaaccaacg ccacctcgat caccggcttg 1440
cagacctccc ttccggccgg tacttacgca gacgagttga gcggactgct taacggcggc 1500
agtattacgg taagttccac gggagccgtc tccgccttta acctgcctgc aggggccgtc 1560
gccgtatggc aatttgcagc gaacgagtcc gcaccggtga ttggcaatat cggaacggac 1620
ctgggcaaag caggcacgac gctgacgatt gatggccgag gctttggcag tacgaaaggg 1680
acggtcttct ttggcacgac ggcagtcacc ggcagccaga ttttgcagtg ggaagattcc 1740
cagatccaag tgacggttcc ggcaatcacc ccagggcaat ataacatcaa agtccgcagc 1800
gcttcctccg tagacagcaa cagctacagc agctttaacc tgctgacagg ggatcaagtc 1860
tcggtccgtt tcatcatcaa caatgccaac acacaattag gcgaaaatgt ctaccttaca 1920
ggcaacatcg ccgaacttgg caactgggat ccgaacaaag cgatgggtcc gttctttaat 1980
cagatcatca ggacgtatcc gacctggtac tacgatgtta gcgtgccagc cggaacaacc 2040
ttggagttca agtttattaa aaagaacggg acagccgtga cctgggaagg cggcgcgaac 2100
cacagcttca caacgccaac gagcggtgtc ggcacggttg aggtcaactg gcagccataa 2160

<210> 578
<211> 719
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<221> SIGNAL
<222> (1)...(34)

<221> DOMAIN
<222> (443)...(527)
<223> Alpha amylase, C-terminal all-beta domain

<221> DOMAIN
<222> (53)...(431)
<223> Alpha amylase, catalytic domain

<221> DOMAIN
<222> (532)...(615)
<223> IPT/TIG domain

<221> DOMAIN
<222> (619)...(715)
<223> Starch binding domain

<400> 578
Met Lys Pro Phe Leu Lys Lys Ser Ile Ile Thr Leu Leu Ala Ser Thr
 1   5   10   15
Cys Leu Phe Thr Ala Trp Leu Ile Pro Ser Ile Ala Val Pro Thr Val
            20   25   30
Ser Ala Ala Pro Asp Thr Ser Val Leu Asn Lys Gln Gly Phe Ser Thr
        35   40   45
Asp Val Ile Tyr Gln Ile Val Thr Asp Arg Phe Ala Asp Gly Asp Pro
    50   55   60
Ser Asn Asn Pro Ser Gly Ala Ala Tyr Ser Pro Gly Cys Thr Asn Leu
65   70   75   80
Lys Leu Tyr Cys Gly Gly Asp Trp Arg Gly Ile Ile Asn Lys Ile Asn
                85   90   95
Glu Gly Tyr Phe Ala Gly Met Gly Val Ser Ala Leu Trp Ile Ser Gln
            100   105   110
Pro Val Glu Asn Ile Tyr Ser Val Ile Asp Tyr Ser Gly Val Asn Ser
        115   120   125

```
Thr Ser Tyr His Gly Tyr Trp Ala Arg Asp Phe Lys Arg Thr Asn Pro
    130                 135                 140
Ala Phe Gly Ser Met Ser Asp Phe Gln Glu Leu Ile Asn Thr Ala His
145                 150                 155                 160
Ala His Asn Ile Lys Ile Ile Ile Asp Phe Ala Pro Asn His Thr Ser
                165                 170                 175
Pro Ala Ser Glu Thr Gln Pro Ser Phe Ala Glu Asn Gly Arg Leu Tyr
            180                 185                 190
Asp Asn Gly Asn Leu Ile Ala Gly Tyr Thr Gly Asp Thr Asn Gly Ile
        195                 200                 205
Phe His His Asn Gln Gly Thr Asn Phe Ser Ser Leu Glu Asp Gly Ile
    210                 215                 220
Tyr Arg Asn Leu Tyr Asp Leu Ala Asp Ile Asn His His Asn Asn Val
225                 230                 235                 240
Thr Asp Thr Tyr Phe Lys Asp Ala Ile Lys Leu Trp Leu Asn Met Gly
                245                 250                 255
Ile Asp Gly Ile Arg Val Asp Ala Val Lys His Met Pro Glu Gly Trp
            260                 265                 270
Gln Lys Asn Trp Val Ala Ser Ile Ala Gly Tyr Lys Pro Val Phe Thr
        275                 280                 285
Phe Gly Glu Trp Phe Leu Gly Val Gly Glu Asn Asp Pro Asn Asn Ile
    290                 295                 300
Arg Phe Ala Asn Glu Ser Gly Met Ser Leu Leu Asp Phe Gln Phe Gly
305                 310                 315                 320
Gln Lys Val Arg Gln Val Phe Arg Asp Asn Ser Asp Asn Met Tyr Gly
                325                 330                 335
Leu Asn Asn Met Leu Thr Thr Thr Ala Ala Asn Tyr Lys His Ile His
            340                 345                 350
Asp Gln Val Thr Phe Ile Asp Asn His Asp Met Asp Arg Phe Lys Leu
        355                 360                 365
Gly Asn Asn Asp Arg Arg Leu Glu Gln Ala Leu Ala Phe Thr Leu Met
    370                 375                 380
Ser Arg Gly Val Pro Ala Ile Tyr Tyr Gly Thr Glu Gln Tyr Met Thr
385                 390                 395                 400
Gly Asn Gly Asp Pro Ala Asn Arg Gly Phe Met Ser Gly Phe Ser Thr
                405                 410                 415
Ser Thr Thr Ala Tyr Gln Val Ile Gly Lys Leu Ala Pro Leu Arg Lys
            420                 425                 430
Ser Asn Pro Ala Leu Ala Tyr Gly Ser Thr Gln Glu Arg Trp Ile Asn
        435                 440                 445
Asn Asp Val Phe Ile Phe Glu Arg Lys Phe Gly Ser Ser Val Ala Leu
    450                 455                 460
Val Ala Ile Asn Arg Ser Thr Thr Asn Ala Thr Ser Ile Thr Gly Leu
465                 470                 475                 480
Gln Thr Ser Leu Pro Ala Gly Thr Tyr Ala Asp Glu Leu Ser Gly Leu
                485                 490                 495
Leu Asn Gly Gly Ser Ile Thr Val Ser Ser Thr Gly Ala Val Ser Ala
            500                 505                 510
Phe Asn Leu Pro Ala Gly Ala Val Ala Val Trp Gln Phe Ala Ala Asn
        515                 520                 525
Glu Ser Ala Pro Val Ile Gly Asn Ile Gly Thr Asp Leu Gly Lys Ala
    530                 535                 540
Gly Thr Thr Leu Thr Ile Asp Gly Arg Gly Phe Gly Ser Thr Lys Gly
545                 550                 555                 560
Thr Val Phe Phe Gly Thr Thr Ala Val Thr Gly Ser Gln Ile Leu Gln
                565                 570                 575
Trp Glu Asp Ser Gln Ile Gln Val Thr Val Pro Ala Ile Thr Pro Gly
            580                 585                 590
Gln Tyr Asn Ile Lys Val Arg Ser Ala Ser Ser Val Asp Ser Asn Ser
        595                 600                 605
Tyr Ser Ser Phe Asn Leu Leu Thr Gly Asp Gln Val Ser Val Arg Phe
    610                 615                 620
Ile Ile Asn Asn Ala Asn Thr Gln Leu Gly Glu Asn Val Tyr Leu Thr
625                 630                 635                 640
Gly Asn Ile Ala Glu Leu Gly Asn Trp Asp Pro Asn Lys Ala Met Gly
                645                 650                 655
Pro Phe Phe Asn Gln Ile Ile Arg Thr Tyr Pro Thr Trp Tyr Tyr Asp
            660                 665                 670
```

Val Ser Val Pro Ala Gly Thr Thr Leu Glu Phe Lys Phe Ile Lys Lys
675 680 685
Asn Gly Thr Ala Val Thr Trp Glu Gly Gly Ala Asn His Ser Phe Thr
690 695 700
Thr Pro Thr Ser Gly Val Gly Thr Val Glu Val Asn Trp Gln Pro
705 710 715

<210> 579
<211> 5007
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 579
gtgtttaaaa ggagagcttt aggatttttg ctggcttttc ttttagtttt tacagcagta 60
tttggctcaa tgcctatgga atttgcaaag gctgagacag atacagcgcc agcgatagcc 120
aatgttgttg gcaattttca atccaaactt ggagattctg actggaatat aaacagtgac 180
aaaacgataa tgacatataa aggtaatggc ttttatgaat ttactacccc agttgcgtta 240
cctgcaggtg attatgagta taaagttgct cttaatcatt catgggaagg tggaggagtt 300
ccttcacaag gtaatttaag ctttcatctt gattcagatt ctgtagtgac tttttattac 360
aattacaaca cttcaagtat taccgattcc acaaaatata caccaattcc cgaagacaag 420
cttccaagac ttgtagggac tatacagcct gcaataggag caggtgatga ttggaaacct 480
gaaacatcga ctgctataat gagagactat aagtttaata atgtttacga atacactgca 540
aatgttccaa aagggaatta tgagtttaaa gtgactttag ggccctcatg ggatataaat 600
tatggcttaa atggtgaaca aaatgggcca aatattcctt tgaatgtagc ctatgatact 660
aagattacat tttactatga ttcggtttca cataatatat ggacagatta taatccacct 720
cttacaggac ctgataataa catatattat gacgatttaa gacatgacac ccatgaccca 780
ttcttccgct cgcctttcgg tgcaataaaa acgggagata ctgtgacatt aagaatacaa 840
gcaaaaaatc atgacattga gtcagctaaa atttcttatt gggatgatat taaaaaaaca 900
agaatagaag tccctatgta cagaattggt caaagtcctg acgggaaata tgaatactgg 960
gaagttaagt taagctttga ccatcccaca agaatttggt attactttat acttaaagac 1020
gggaccaaaa ctgcttatta cggagataac gatgaacaat taggtggcgt aggtaaagcc 1080
acagatacag aaaataagga cttcgaactt actgtctacg ataaaaactt agacacccct 1140
gattggatga aagggtcagt aatgtatcaa atattccctg ataggttctt taatggggat 1200
tcttcaaatg accatctaaa gaaatacagc agaagttttg atcctgttga atatcatagc 1260
aactggtacg agcttccaga taatccgaat gataaaaata aactaggata tacaggggat 1320
ggcatatggt caaatgactt ctttggcggt gatttaaaag gtatagatga taaattggat 1380
tatttaaaaa gccttggaat atcagtcatt tatctcaatc caatttttca atcaccttct 1440
aatcacaggt acgatacaac cgattacaca aaaatagacg aactattagg agatttatct 1500
acctttaaga agctcatgga ggatgcccat gcaaaaggga ttaaggtaat acttgatggc 1560
gtcttcaatc atacaagtga tgacagcatt tattttgata gatatggtaa atacttaaac 1620
acgggagttt taggtgctta tcaagcatgg aaacagggag atcagtcaaa gtctccatat 1680
ggtgactggt atgagataaa gcctgacggt acctatgaag gatggtgggg atttgacagt 1740
ttacctgtga taaggcagat aaacggcagt gaatataatg tcaaaagttg ggcggatttt 1800
attataaata atcctaatgc aatatctaag tattggttaa atcctgatgg ggacaaaaat 1860
gtaggtgcag atggttggag attagatgtt gcgaatgaag ttgctcacga tttttgggtt 1920
cattttagag gtgcaattaa tactgtgaaa ccgaatgctc cgatggttgc agaaaactgg 1980
aatgacgctt cactggatct gcttggagat tcttttaact ctgttatgaa ttatctcttt 2040
agaaatgcag taattgactt tatattggat aaatcatttg atgacggaaa tgtggttcac 2100
aatcctatag atgcagcaaa actcgaccaa aggcttatga gcatatatga gagatatcct 2160
cttccagtgt tttattctac tatgaatctt ttaggttctc atgacaccat gaggatattg 2220
acagtatttg gatacaactc tgcggatgaa aatcaaaatt ctcaagcagc aaaagacctt 2280
gcggttaaaa gacttaaact tgctgcaata ctgcaaatgg gttatccagg catgccttct 2340
atttattacg gagatgaagc aggacaatct ggcggaaaag acccggacaa cagaaggaca 2400
ttcccttggg gaagagaaga tactgatttg caaacttttt tcaagaaagt gtaaatata 2460
agaaatgaga atcaagtttt gaaaacagga gaccttgaga cactttatgc aaatggcgat 2520
gtttatgcct ttggaagaag aatcataaat ggaaaagata cttttgggaa atcgtatcct 2580
gacagtgtag ctattgtcgt aataaataaa ggtgacgcaa aacaggtttc tatagatacc 2640
actaaattta taagagatgg agttgctttt acagatgcct taagtggcaa gacatacacg 2700
gttcaggatg gtaaaattgt tgtagaagtt ggatcaatgg atggagctat acttatatca 2760
gatacaggac aaaatttgac agcacctcag ccaataacag accttaaagc agtctcaggg 2820
aatggcaaag tagacctttc ctggagtgta gtagataaag cagtaagcta taatatttac 2880
cgttctacag ttaaaggtgg gttatatgaa aaaatagctt caaatgttac gcaaattacc 2940
tatactgata cagaggttac caatggtcta aagtatgtgt atgctgtaac cgctgtagat 3000
aatgatggaa atgagagtgc tttaagcaat gaagttgagg catatccagc atttcctatt 3060
ggctgggcgg gaaatatgaa tcaagtaaat acacatgtaa taggtgtaaa taatccagtt 3120

gaagtgtatg ctgaagtgtg ggctcaaggg cttacagata aacctggcca aggagaaaat 3180
atgattgccc agttaggata taggtatatt ggagatactg taggagatgc agtttacaat 3240
gccgtgtaca ataaggtaga aggtgttgaa ataagtaagg actggacatg ggttgatgca 3300
caatatgtag gtgattctgg aaataatgat aaatacatgg ctaaatttgt acctgatatg 3360
gtaggcacat gggaatacat tatgagattt tcttctaacc aaggccatga ttggacgtat 3420
acaaaaggac cagatgggaa aacagatgaa gcaaaacagt ttactgtcgt gccatcaaat 3480
gatgtagaaa cacctacagc tccagtctta caacaaccag gaattgaatc ctcaagagtt 3540
acacttaact ggagtccgtc agctgatgat gttgctattt tcggctacga aatctataag 3600
tcttcaagtg agacaggacc atttataaag attgcaactg tgtctgacag tgtgtataac 3660
tacgtagata cagatgtagt aaatggcaat gtgtattatt ataaagttgt ggcagttgat 3720
acctcctata ataggacagc atcaaataca gtaaaagcga cacctgatat aatacctata 3780
aaagtgacat ttaatgtcac aataccagat tatactcctg atgatggtgt aaatattgct 3840
ggaaacttcc cggatgcttt ctggaatcca aatgctaatc aaatgacaaa agctggttct 3900
aacacctata gtattacact tactttaaat gagggtacac aaattgaata taaatatgca 3960
agggggagct gggataaagt agaaaaaggt gaatacggca atgagattga caacagaaaa 4020
ataactgttg tcaatcaggg ttcaaataca atggttgtaa atgatacagt gcaaagatgg 4080
agagatgtgc caatatatat ctattctcca aaagacaaga ctattgttga tgcaaatacg 4140
agtgaaatag agattaaagg caatacttac aaaggtgcga aagtaactat aaatgatgaa 4200
tcttttgtgc aacaagaaaa tggagtattt acaaaagtag ttcccttgga atatggtgta 4260
aatactatca aaatacatgt agagccaagt ggtgacaaga ataatgaact tacaaaagat 4320
ataacaataa ctgttacaag agagaagccc ggccaggaga cagaacctac tcctactcca 4380
ccaacagaaa caacaaaacc atctcaagaa gtatcacaag gcaaaatagt tgtagaaaat 4440
aacacaacaa cacttacaat agacgaaaac aaagtagcaa aagacataaa agacacttca 4500
aagaaagaaa tacgatttga cttaacaaat ataggaacta cacctcaaaa agcattagaa 4560
atacctgtaa ctgtacttaa tttaatagca gagaacaaca aaaacgtagt agtaaaatca 4620
gacgaagtag ctcttcaatt tgatgcaaaa actttagcag tatcacaaga agcaattgac 4680
ctcataaaca aagcaggaac aataaaactt aacatacaca acaaaggcaa aaaggaagca 4740
tcatctttcg agccaataac aagtgcctat gacataacaa taaaagcagg agacaaagac 4800
ataaaaatag ggtcaccagt aaaaatgaca tttaacataa aaggtggaaa agacataaga 4860
aaagcaggag tatactactt aaacgagaag acaaatcaat gggaatacgt aggaggaaaa 4920
attgacaaag gaacaaacac aataaccttt gaagccaaac acttctcaac ctatggagtc 4980
tttgaataca acaaagaatt caaaaag 5007

<210> 580
<211> 1669
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<221> SIGNAL
<222> (1)...(29)

<221> DOMAIN
<222> (390)...(821)
<223> Alpha amylase, catalytic domain

<221> DOMAIN
<222> (833)...(924)
<223> Alpha amylase, C-terminal all-beta domain

<221> DOMAIN
<222> (929)...(1009)
<223> Fibronectin type III domain

<221> DOMAIN
<222> (1165)...(1248)
<223> Fibronectin type III domain

<400> 580
Met Phe Lys Arg Arg Ala Leu Gly Phe Leu Leu Ala Phe Leu Leu Val
 1   5   10   15
Phe Thr Ala Val Phe Gly Ser Met Pro Met Glu Phe Ala Lys Ala Glu
            20   25   30
Thr Asp Thr Ala Pro Ala Ile Ala Asn Val Val Gly Asn Phe Gln Ser

```
        35                  40                  45
Lys Leu Gly Asp Ser Asp Trp Asn Ile Asn Ser Asp Lys Thr Ile Met
    50                  55                  60
Thr Tyr Lys Gly Asn Gly Phe Tyr Glu Phe Thr Thr Pro Val Ala Leu
65                  70                  75                  80
Pro Ala Gly Asp Tyr Glu Tyr Lys Val Ala Leu Asn His Ser Trp Glu
                85                  90                  95
Gly Gly Gly Val Pro Ser Gln Gly Asn Leu Ser Phe His Leu Asp Ser
            100                 105                 110
Asp Ser Val Val Thr Phe Tyr Tyr Asn Tyr Asn Thr Ser Ser Ile Thr
        115                 120                 125
Asp Ser Thr Lys Tyr Thr Pro Ile Pro Glu Asp Lys Leu Pro Arg Leu
    130                 135                 140
Val Gly Thr Ile Gln Pro Ala Ile Gly Ala Gly Asp Asp Trp Lys Pro
145                 150                 155                 160
Glu Thr Ser Thr Ala Ile Met Arg Asp Tyr Lys Phe Asn Asn Val Tyr
                165                 170                 175
Glu Tyr Thr Ala Asn Val Pro Lys Gly Asn Tyr Glu Phe Lys Val Thr
            180                 185                 190
Leu Gly Pro Ser Trp Asp Ile Asn Tyr Gly Leu Asn Gly Glu Gln Asn
        195                 200                 205
Gly Pro Asn Ile Pro Leu Asn Val Ala Tyr Asp Thr Lys Ile Thr Phe
    210                 215                 220
Tyr Tyr Asp Ser Val Ser His Asn Ile Trp Thr Asp Tyr Asn Pro Pro
225                 230                 235                 240
Leu Thr Gly Pro Asp Asn Asn Ile Tyr Tyr Asp Asp Leu Arg His Asp
                245                 250                 255
Thr His Asp Pro Phe Phe Arg Ser Pro Phe Gly Ala Ile Lys Thr Gly
            260                 265                 270
Asp Thr Val Thr Leu Arg Ile Gln Ala Lys Asn His Asp Ile Glu Ser
        275                 280                 285
Ala Lys Ile Ser Tyr Trp Asp Asp Ile Lys Lys Thr Arg Ile Glu Val
    290                 295                 300
Pro Met Tyr Arg Ile Gly Gln Ser Pro Asp Gly Lys Tyr Glu Tyr Trp
305                 310                 315                 320
Glu Val Lys Leu Ser Phe Asp His Pro Thr Arg Ile Trp Tyr Tyr Phe
                325                 330                 335
Ile Leu Lys Asp Gly Thr Lys Thr Ala Tyr Tyr Gly Asp Asn Asp Glu
            340                 345                 350
Gln Leu Gly Gly Val Gly Lys Ala Thr Asp Thr Glu Asn Lys Asp Phe
        355                 360                 365
Glu Leu Thr Val Tyr Asp Lys Asn Leu Asp Thr Pro Asp Trp Met Lys
    370                 375                 380
Gly Ser Val Met Tyr Gln Ile Phe Pro Asp Arg Phe Phe Asn Gly Asp
385                 390                 395                 400
Ser Ser Asn Asp His Leu Lys Lys Tyr Ser Arg Ser Phe Asp Pro Val
                405                 410                 415
Glu Tyr His Ser Asn Trp Tyr Glu Leu Pro Asp Asn Pro Asn Asp Lys
            420                 425                 430
Asn Lys Leu Gly Tyr Thr Gly Asp Gly Ile Trp Ser Asn Asp Phe Phe
        435                 440                 445
Gly Gly Asp Leu Lys Gly Ile Asp Asp Lys Leu Asp Tyr Leu Lys Ser
    450                 455                 460
Leu Gly Ile Ser Val Ile Tyr Leu Asn Pro Ile Phe Gln Ser Pro Ser
465                 470                 475                 480
Asn His Arg Tyr Asp Thr Thr Asp Tyr Thr Lys Ile Asp Glu Leu Leu
                485                 490                 495
Gly Asp Leu Ser Thr Phe Lys Lys Leu Met Glu Asp Ala His Ala Lys
            500                 505                 510
Gly Ile Lys Val Ile Leu Asp Gly Val Phe Asn His Thr Ser Asp Asp
        515                 520                 525
Ser Ile Tyr Phe Asp Arg Tyr Gly Lys Tyr Leu Asn Thr Gly Val Leu
    530                 535                 540
Gly Ala Tyr Gln Ala Trp Lys Gln Gly Asp Gln Ser Lys Ser Pro Tyr
545                 550                 555                 560
Gly Asp Trp Tyr Glu Ile Lys Pro Asp Gly Thr Tyr Glu Gly Trp Trp
                565                 570                 575
Gly Phe Asp Ser Leu Pro Val Ile Arg Gln Ile Asn Gly Ser Glu Tyr
```

```
            580                 585                 590
Asn Val Lys Ser Trp Ala Asp Phe Ile Ile Asn Asn Pro Asn Ala Ile
        595                 600                 605
Ser Lys Tyr Trp Leu Asn Pro Asp Gly Asp Lys Asn Val Gly Ala Asp
    610                 615                 620
Gly Trp Arg Leu Asp Val Ala Asn Glu Val Ala His Asp Phe Trp Val
625                 630                 635                 640
His Phe Arg Gly Ala Ile Asn Thr Val Lys Pro Asn Ala Pro Met Val
                645                 650                 655
Ala Glu Asn Trp Asn Asp Ala Ser Leu Asp Leu Leu Gly Asp Ser Phe
            660                 665                 670
Asn Ser Val Met Asn Tyr Leu Phe Arg Asn Ala Val Ile Asp Phe Ile
        675                 680                 685
Leu Asp Lys Ser Phe Asp Asp Gly Asn Val Val His Asn Pro Ile Asp
    690                 695                 700
Ala Ala Lys Leu Asp Gln Arg Leu Met Ser Ile Tyr Glu Arg Tyr Pro
705                 710                 715                 720
Leu Pro Val Phe Tyr Ser Thr Met Asn Leu Leu Gly Ser His Asp Thr
                725                 730                 735
Met Arg Ile Leu Thr Val Phe Gly Tyr Asn Ser Ala Asp Glu Asn Gln
            740                 745                 750
Asn Ser Gln Ala Ala Lys Asp Leu Ala Val Lys Arg Leu Lys Leu Ala
        755                 760                 765
Ala Ile Leu Gln Met Gly Tyr Pro Gly Met Pro Ser Ile Tyr Tyr Gly
    770                 775                 780
Asp Glu Ala Gly Gln Ser Gly Gly Lys Asp Pro Asp Asn Arg Arg Thr
785                 790                 795                 800
Phe Pro Trp Gly Arg Glu Asp Thr Asp Leu Gln Thr Phe Phe Lys Lys
                805                 810                 815
Val Val Asn Ile Arg Asn Glu Asn Gln Val Leu Lys Thr Gly Asp Leu
            820                 825                 830
Glu Thr Leu Tyr Ala Asn Gly Asp Val Tyr Ala Phe Gly Arg Arg Ile
        835                 840                 845
Ile Asn Gly Lys Asp Thr Phe Gly Lys Ser Tyr Pro Asp Ser Val Ala
    850                 855                 860
Ile Val Val Ile Asn Lys Gly Asp Ala Lys Gln Val Ser Ile Asp Thr
865                 870                 875                 880
Thr Lys Phe Ile Arg Asp Gly Val Ala Phe Thr Asp Ala Leu Ser Gly
                885                 890                 895
Lys Thr Tyr Thr Val Gln Asp Gly Lys Ile Val Val Glu Val Gly Ser
            900                 905                 910
Met Asp Gly Ala Ile Leu Ile Ser Asp Thr Gly Gln Asn Leu Thr Ala
        915                 920                 925
Pro Gln Pro Ile Thr Asp Leu Lys Ala Val Ser Gly Asn Gly Lys Val
    930                 935                 940
Asp Leu Ser Trp Ser Val Val Asp Lys Ala Val Ser Tyr Asn Ile Tyr
945                 950                 955                 960
Arg Ser Thr Val Lys Gly Gly Leu Tyr Glu Lys Ile Ala Ser Asn Val
                965                 970                 975
Thr Gln Ile Thr Tyr Thr Asp Thr Glu Val Thr Asn Gly Leu Lys Tyr
            980                 985                 990
Val Tyr Ala Val Thr Ala Val Asp Asn Asp Gly Asn Glu Ser Ala Leu
        995                 1000                1005
Ser Asn Glu Val Glu Ala Tyr Pro Ala Phe Pro Ile Gly Trp Ala Gly
    1010                1015                1020
Asn Met Asn Gln Val Asn Thr His Val Ile Gly Val Asn Asn Pro Val
1025                1030                1035                1040
Glu Val Tyr Ala Glu Val Trp Ala Gln Gly Leu Thr Asp Lys Pro Gly
                1045                1050                1055
Gln Gly Glu Asn Met Ile Ala Gln Leu Gly Tyr Arg Tyr Ile Gly Asp
            1060                1065                1070
Thr Val Gly Asp Ala Val Tyr Asn Ala Val Tyr Asn Lys Val Glu Gly
        1075                1080                1085
Val Glu Ile Ser Lys Asp Trp Thr Trp Val Asp Ala Gln Tyr Val Gly
    1090                1095                1100
Asp Ser Gly Asn Asn Asp Lys Tyr Met Ala Lys Phe Val Pro Asp Met
1105                1110                1115                1120
Val Gly Thr Trp Glu Tyr Ile Met Arg Phe Ser Ser Asn Gln Gly His
```

1125 1130 1135
Asp Trp Thr Tyr Thr Lys Gly Pro Asp Gly Lys Thr Asp Glu Ala Lys
1140 1145 1150
Gln Phe Thr Val Val Pro Ser Asn Asp Val Glu Thr Pro Thr Ala Pro
1155 1160 1165
Val Leu Gln Gln Pro Gly Ile Glu Ser Ser Arg Val Thr Leu Asn Trp
1170 1175 1180
Ser Pro Ser Ala Asp Asp Val Ala Ile Phe Gly Tyr Glu Ile Tyr Lys
1185 1190 1195 1200
Ser Ser Ser Glu Thr Gly Pro Phe Ile Lys Ile Ala Thr Val Ser Asp
1205 1210 1215
Ser Val Tyr Asn Tyr Val Asp Thr Asp Val Val Asn Gly Asn Val Tyr
1220 1225 1230
Tyr Tyr Lys Val Val Ala Val Asp Thr Ser Tyr Asn Arg Thr Ala Ser
1235 1240 1245
Asn Thr Val Lys Ala Thr Pro Asp Ile Ile Pro Ile Lys Val Thr Phe
1250 1255 1260
Asn Val Thr Ile Pro Asp Tyr Thr Pro Asp Asp Gly Val Asn Ile Ala
1265 1270 1275 1280
Gly Asn Phe Pro Asp Ala Phe Trp Asn Pro Asn Ala Asn Gln Met Thr
1285 1290 1295
Lys Ala Gly Ser Asn Thr Tyr Ser Ile Thr Leu Thr Leu Asn Glu Gly
1300 1305 1310
Thr Gln Ile Glu Tyr Lys Tyr Ala Arg Gly Ser Trp Asp Lys Val Glu
1315 1320 1325
Lys Gly Glu Tyr Gly Asn Glu Ile Asp Asn Arg Lys Ile Thr Val Val
1330 1335 1340
Asn Gln Gly Ser Asn Thr Met Val Val Asn Asp Thr Val Gln Arg Trp
1345 1350 1355 1360
Arg Asp Val Pro Ile Tyr Ile Tyr Ser Pro Lys Asp Lys Thr Ile Val
1365 1370 1375
Asp Ala Asn Thr Ser Glu Ile Glu Ile Lys Gly Asn Thr Tyr Lys Gly
1380 1385 1390
Ala Lys Val Thr Ile Asn Asp Glu Ser Phe Val Gln Gln Glu Asn Gly
1395 1400 1405
Val Phe Thr Lys Val Val Pro Leu Glu Tyr Gly Val Asn Thr Ile Lys
1410 1415 1420
Ile His Val Glu Pro Ser Gly Asp Lys Asn Asn Glu Leu Thr Lys Asp
1425 1430 1435 1440
Ile Thr Ile Thr Val Thr Arg Glu Lys Pro Gly Gln Glu Thr Glu Pro
1445 1450 1455
Thr Pro Thr Pro Pro Thr Glu Thr Thr Lys Pro Ser Gln Glu Val Ser
1460 1465 1470
Gln Gly Lys Ile Val Val Glu Asn Asn Thr Thr Thr Leu Thr Ile Asp
1475 1480 1485
Glu Asn Lys Val Ala Lys Asp Ile Lys Asp Thr Ser Lys Lys Glu Ile
1490 1495 1500
Arg Phe Asp Leu Thr Asn Ile Gly Thr Thr Pro Gln Lys Ala Leu Glu
1505 1510 1515 1520
Ile Pro Val Thr Val Leu Asn Leu Ile Ala Glu Asn Asn Lys Asn Val
1525 1530 1535
Val Val Lys Ser Asp Glu Val Ala Leu Gln Phe Asp Ala Lys Thr Leu
1540 1545 1550
Ala Val Ser Gln Glu Ala Ile Asp Leu Ile Asn Lys Ala Gly Thr Ile
1555 1560 1565
Lys Leu Asn Ile His Asn Lys Gly Lys Lys Glu Ala Ser Ser Phe Glu
1570 1575 1580
Pro Ile Thr Ser Ala Tyr Asp Ile Thr Ile Lys Ala Gly Asp Lys Asp
1585 1590 1595 1600
Ile Lys Ile Gly Ser Pro Val Lys Met Thr Phe Asn Ile Lys Gly Gly
1605 1610 1615
Lys Asp Ile Arg Lys Ala Gly Val Tyr Tyr Leu Asn Glu Lys Thr Asn
1620 1625 1630
Gln Trp Glu Tyr Val Gly Gly Lys Ile Asp Lys Gly Thr Asn Thr Ile
1635 1640 1645
Thr Phe Glu Ala Lys His Phe Ser Thr Tyr Gly Val Phe Glu Tyr Asn
1650 1655 1660
Lys Glu Phe Lys Lys

1665

<210> 581
<211> 2142
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 581
atgaaaaagt tttacaaact gacaaccgct cttgccctta gcttgtccct agcgttgagc 60
cttctcggcc ctgctcatgc ggccccggat acttcggtgt ccaataagca aaattttagc 120
acggatgtca tttatcaaat ctttacggat cgtttctccg acggcaatgc tgccaacaac 180
ccgaccggac ctgcctttga cggtacctgt accaatctga ggctgtactg cggcggcgac 240
tggcaaggga tcattaacaa aatcaacgac ggctacctga ccggcatggg cgttaccgcc 300
atctggatct cccagcctgt cgaaaacatt tacagcgtta tcaactattc cggcgtgagc 360
aacaccgctt atcatggcta ttgggccaga gacttcaaga aaacgaatcc ggcttacggc 420
acgatcgccg atttccagaa tctcatcgca gccgctcatg cgaagaacat caaagtcatc 480
atcgacttcg caccgaacca tacctctccc gcttcctcgg accagccttc attcgcggag 540
aacggcaagc tgtacaacaa cggcaccctg ctcggcggat acacgggcga tacacagaat 600
ttgttccatc ataacggggg caccgatttt tcaacgacag agaacggaat ttacaaaaac 660
ctgtacgacc ttgctgacct gaaccacaac aacggcaccg tggataccta cctgaaggat 720
gccatcaaaa tgtggctcga tctgggcatt gacggcattc gaatggacgc ggtgaaacat 780
atgccgtttg gctggcagaa gagtttcatg gccaccgtta acaactataa gccggtcttc 840
accttcggcg aatggttcct cggcgtgaac gaagtgagtg cggagaacca taagttcgcc 900
aacgaatccg gcatgagcct gctggacttc cgtttcgcgc agaaggtacg ccaggtattc 960
agggacaata ccgacaatat gtacggtctg aagtccatgc tggagggctc ggcaaccgac 1020
tatgcccaag tggaggatca ggtcacattc attgataacc atgacatgga acgtttccac 1080
gataacagcg ccaaccgccg gaagctggaa caggccttgg ccttcacgct cacctcacgc 1140
ggggtgccag ccatctatta cggaaccgaa caatacatgt ccggcgggaa cgatcctaac 1200
aaccgggcac ggattccttc attctctacg acgactaccg cttaccaagt aattaaaaag 1260
ctcgcgccat tgcgtaaatc caatccggcc atcgcctatg gaacgacgca ggagcgctgg 1320
atcaacaatg atgtgctgat ctatgaacgc aaattcggca gcaacgtagc gcttatcgct 1380
gtaaaccgca acgtgaacac gccggcttcc attaccggat tggcgacttc attgcctgcc 1440
ggcagctatg ccgacgtgct gggcggtctc ctgaacggta acagcctgac cgttggcaca 1500
ggcggagcag cttccaactt tacgctggct gccggcggta cggccgtatg gcaatatacc 1560
acggccgtga cttcgccaac gatcgggcac gtaggcccga tgatggccaa accgggcgca 1620
accgtcacca ttgacggccg cggcttcggc gcaaccaaag gcacggtata cttcggctca 1680
acggccgtaa cgggcgctaa cattacggca tgggaggaca cgcagatcaa agtgaagatc 1740
cctgcggttg ccggtggcgt atataacatt aaagttgcca atatcgctgg cacttcaagc 1800
aatgtgcacg acaacttcga agtactgacc ggggatcaag tcagtgtccg gtttgtggtg 1860
aacaatgcca ccaccgcact cggacaaaac gtgtacctgg caggcagcgt aagcgagctg 1920
ggcaattggg acccggccaa agccatcggt ccgttgtaca accaggttat ctaccaatac 1980
ccaacttggt actacgacgt cagcgtaccg gccggtaaaa ccatcgaatt caaattcctg 2040
aaaaagcagg gctcgacggt aacgtgggaa ggcggatcga accatacctt taccgcacca 2100
acgagcggca cggccacgat taatgtaaac tggcagccgt aa 2142

<210> 582
<211> 713
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<221> SIGNAL
<222> (1)...(27)

<221> DOMAIN
<222> (46)...(425)
<223> Alpha amylase, catalytic domain

<221> DOMAIN
<222> (526)...(609)
<223> IPT/TIG domain

<221> DOMAIN

<222> (613)...(709)
<223> Starch binding domain

<221> DOMAIN
<222> (437)...(521)
<223> Alpha amylase, C-terminal all-beta domain

<400> 582

```
Met Lys Lys Phe Tyr Lys Leu Thr Thr Ala Leu Ala Leu Ser Leu Ser
 1               5                  10                  15
Leu Ala Leu Ser Leu Leu Gly Pro Ala His Ala Ala Pro Asp Thr Ser
            20                  25                  30
Val Ser Asn Lys Gln Asn Phe Ser Thr Asp Val Ile Tyr Gln Ile Phe
        35                  40                  45
Thr Asp Arg Phe Ser Asp Gly Asn Ala Ala Asn Asn Pro Thr Gly Pro
    50                  55                  60
Ala Phe Asp Gly Thr Cys Thr Asn Leu Arg Leu Tyr Cys Gly Gly Asp
65                  70                  75                  80
Trp Gln Gly Ile Ile Asn Lys Ile Asn Asp Gly Tyr Leu Thr Gly Met
                85                  90                  95
Gly Val Thr Ala Ile Trp Ile Ser Gln Pro Val Glu Asn Ile Tyr Ser
            100                 105                 110
Val Ile Asn Tyr Ser Gly Val Ser Asn Thr Ala Tyr His Gly Tyr Trp
        115                 120                 125
Ala Arg Asp Phe Lys Lys Thr Asn Pro Ala Tyr Gly Thr Ile Ala Asp
    130                 135                 140
Phe Gln Asn Leu Ile Ala Ala Ala His Ala Lys Asn Ile Lys Val Ile
145                 150                 155                 160
Ile Asp Phe Ala Pro Asn His Thr Ser Pro Ala Ser Ser Asp Gln Pro
                165                 170                 175
Ser Phe Ala Glu Asn Gly Lys Leu Tyr Asn Asn Gly Thr Leu Leu Gly
            180                 185                 190
Gly Tyr Thr Gly Asp Thr Gln Asn Leu Phe His His Asn Gly Gly Thr
        195                 200                 205
Asp Phe Ser Thr Thr Glu Asn Gly Ile Tyr Lys Asn Leu Tyr Asp Leu
    210                 215                 220
Ala Asp Leu Asn His Asn Asn Gly Thr Val Asp Thr Tyr Leu Lys Asp
225                 230                 235                 240
Ala Ile Lys Met Trp Leu Asp Leu Gly Ile Asp Gly Ile Arg Met Asp
                245                 250                 255
Ala Val Lys His Met Pro Phe Gly Trp Gln Lys Ser Phe Met Ala Thr
            260                 265                 270
Val Asn Asn Tyr Lys Pro Val Phe Thr Phe Gly Glu Trp Phe Leu Gly
        275                 280                 285
Val Asn Glu Val Ser Ala Glu Asn His Lys Phe Ala Asn Glu Ser Gly
    290                 295                 300
Met Ser Leu Leu Asp Phe Arg Phe Ala Gln Lys Val Arg Gln Val Phe
305                 310                 315                 320
Arg Asp Asn Thr Asp Asn Met Tyr Gly Leu Lys Ser Met Leu Glu Gly
                325                 330                 335
Ser Ala Thr Asp Tyr Ala Gln Val Glu Asp Gln Val Thr Phe Ile Asp
            340                 345                 350
Asn His Asp Met Glu Arg Phe His Asp Asn Ser Ala Asn Arg Arg Lys
        355                 360                 365
Leu Glu Gln Ala Leu Ala Phe Thr Leu Thr Ser Arg Gly Val Pro Ala
    370                 375                 380
Ile Tyr Tyr Gly Thr Glu Gln Tyr Met Ser Gly Gly Asn Asp Pro Asn
385                 390                 395                 400
Asn Arg Ala Arg Ile Pro Ser Phe Ser Thr Thr Thr Thr Ala Tyr Gln
                405                 410                 415
Val Ile Lys Lys Leu Ala Pro Leu Arg Lys Ser Asn Pro Ala Ile Ala
            420                 425                 430
Tyr Gly Thr Thr Gln Glu Arg Trp Ile Asn Asn Asp Val Leu Ile Tyr
        435                 440                 445
Glu Arg Lys Phe Gly Ser Asn Val Ala Leu Ile Ala Val Asn Arg Asn
    450                 455                 460
Val Asn Thr Pro Ala Ser Ile Thr Gly Leu Ala Thr Ser Leu Pro Ala
465                 470                 475                 480
```

```
Gly Ser Tyr Ala Asp Val Leu Gly Gly Leu Leu Asn Gly Asn Ser Leu
                485                 490                 495
Thr Val Gly Thr Gly Gly Ala Ala Ser Asn Phe Thr Leu Ala Ala Gly
            500                 505                 510
Gly Thr Ala Val Trp Gln Tyr Thr Thr Ala Val Thr Ser Pro Thr Ile
        515                 520                 525
Gly His Val Gly Pro Met Met Ala Lys Pro Gly Ala Thr Val Thr Ile
    530                 535                 540
Asp Gly Arg Gly Phe Gly Ala Thr Lys Gly Thr Val Tyr Phe Gly Ser
545                 550                 555                 560
Thr Ala Val Thr Gly Ala Asn Ile Thr Ala Trp Glu Asp Thr Gln Ile
                565                 570                 575
Lys Val Lys Ile Pro Ala Val Ala Gly Gly Val Tyr Asn Ile Lys Val
            580                 585                 590
Ala Asn Ile Ala Gly Thr Ser Ser Asn Val His Asp Asn Phe Glu Val
        595                 600                 605
Leu Thr Gly Asp Gln Val Ser Val Arg Phe Val Val Asn Asn Ala Thr
    610                 615                 620
Thr Ala Leu Gly Gln Asn Val Tyr Leu Ala Gly Ser Val Ser Glu Leu
625                 630                 635                 640
Gly Asn Trp Asp Pro Ala Lys Ala Ile Gly Pro Leu Tyr Asn Gln Val
                645                 650                 655
Ile Tyr Gln Tyr Pro Thr Trp Tyr Tyr Asp Val Ser Val Pro Ala Gly
            660                 665                 670
Lys Thr Ile Glu Phe Lys Phe Leu Lys Lys Gln Gly Ser Thr Val Thr
        675                 680                 685
Trp Glu Gly Gly Ser Asn His Thr Phe Thr Ala Pro Thr Ser Gly Thr
    690                 695                 700
Ala Thr Ile Asn Val Asn Trp Gln Pro
705                 710
```

<210> 583
<211> 6618
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 583

```
atgatggctg tgtcaaaaaa agaaatcacc caagacatgt tccagccggc ggtcatcgat      60
aaaagcctgc aagaacaagt ttccggaccg agcatcggct tttggagcga cggtttccgg     120
cgcctgcgca aaaataaagc ggcgatggtg agcgcgttca ttatcctgct catcgtcatc     180
atggcgctcg tcgggccgca catgagcaaa tacggcatcg atgaccaaga cttgaaacgc     240
tcgaacttac cgccgcgcat ttccgcattg gccaacattc attggctgcc ttttgacggc     300
aaaaacgaac aaggcgtcga cgtttacaaa gagcgcggca tcaaacaaga tttgtggttt     360
ggcgcggaca aactcggccg cgacctgtgg acacgggttt ggaaaggcac gcaaatttcc     420
ctcttcatcg ccgtcgtcgc cgctgcggcc gacctgatta tcggcgtcgc ttacgggggg     480
atttccgcct atttcggcgg ccaggtcgac aacatcatgc agaggatact cgaagtgctc     540
aatggcattc ccaacctggt gatcatcata ttgttaatca tcgtcatgca accgggcata     600
cccgcgatca tcctggcgat cgccatgacc ggttggatcg gcatggcgcg gatcatacgc     660
ggagaagttt taaaacttaa aaatttggaa ttcaagctga ccaagaacgg cagctacgcg     720
tccggcattc cgggagacgg ggaaaatttg aagctgacgc tcagccaggc cgccaaggtg     780
aatttttatg taaacgggga cacgcaagaa gcccgaatat cgccgctgcc ggagggatcg     840
accgcgagcg gcattcagat ctacaacccg attttgtccg cggataagtg gccgcgtctg     900
gtcggttcgt tgcagccgct attcggcgag cccgaatggt cgccgggaga ggccaagcag     960
ttttttgtcg attataactt caacaatacc gtgtataagc ttcagcgctc gcttccggca    1020
ggctcgtacg aggcgaaggt cgtgctgggc gcggattggg cggacggcaa caacttcgga    1080
aacggaaacg ccaacgccgc attcagactc ccggacgccg cggatgtgac cttttccttc    1140
gacaattcgg cggctgagag gaaactcacg gtcgattacg cggtcggaga cggcaagtat    1200
gacggcaaaa tcgatgcggg cgcgattgcg ttcgacagcc ggcaggcaac gttcaaaaag    1260
ccgttcggcg ccatcaagca aggcgagcag gacttgacgc tgcgcatcgc cgtcaaacgc    1320
gacgacgtgc agatggcccg ggtggagctg accgacgccc atggcgcggc gaccaactat    1380
gtgatgcgca aaacgacgac agtcggggat cgcgattatt acgaagcggt cattccgaaa    1440
agcgtttttg ccggcattgg catctggggt tataaattcg tgctgatcga cggggcggcg    1500
aaggtcgaat acggcgacga cagtctcagc ggcggctcgg gcgaggcggt ggcggacggc    1560
gcggttccgt ttaacctgac ggtgtacgcg gcggactaca aaacgccgga ttggatgaaa    1620
acggccgtcg tgtatcagat ttttcccgac cgcttctttg acggcaatcc ggacaacaac    1680
```

```
cgcgcgaagc tggccgacgg ttaccggggc gttcgcagcg agaccgggga acaagccggg    1740
cgcatcggca aacaggagct gcaatatttc gacggaggcg tagcgaacga tccggctcct    1800
tcgcaggtat ggggcgattg gaacgacccg ccggagaacc cggaccgcac caagccggaa    1860
aacaagcctt attatccgga cgcgaagtcg gacggcatct ggacgaacga gttttacggc    1920
ggggacattc agggcatcga gcagaagctc gactatctga agcagcttgg cgtcacggcg    1980
atttacctca atcccgtcgc ctgggcggct tccaaccata aatacgatgc gaccgactac    2040
aagcatctcg atccgatgtt cggcgaaccg gtatacaaca cgcccggcga cccgtcttcc    2100
ggattgaatt acgaggagac ccggaaaaag tcggacgaag tgtaccggaa attcgccaag    2160
gcggcagccg aggccgggat tcggatcatc aacgacggcg tgttcaacca tgtcggggac    2220
gactccattt atttcgaccg ttacgagaaa tatccggaaa tcggcgccta cgagtattgg    2280
tcgaaggtgt ggaagacgaa ggccgaacat ccgggcatga cccgcgagga ggccgagcag    2340
cgcgtcatcg cttcgtacac gagccaaatc aatccggcaa ccgggcggca ttaccaatat    2400
ccggacgatt tcggcttcac cacgtggttt acgatcggcg aaaatatggc ggacaatacc    2460
cgttacagct acgacgcctg gtggggattc gactcgctgc ccgccatgga cgccaaggaa    2520
ccgcagccgg gcgataaaga cgcgctcgaa ggcgttcacg aatggaacaa tatcgactac    2580
cgcaatcatg tcatcgggta tgatctcagc ggcaaatcgg aggccgaagc gtccgcagcg    2640
atgcaaaaca cggcgtcgca gcgatggctg tggatgggct cgagcggctg gcgtctcgac    2700
gtggcgccgg acgtgtcggc aggcacgtgg cggaaattcc gcgaggcggt caaatccgcg    2760
gcgggacgaa cggacgcaaa cggcaagccg atcgacgatc cggtcattct cggcgaagag    2820
tggggcgtgg caacgcatta tttgctcggg gaccagtttg actcgatcat gaactaccgg    2880
ttccgcgaag ccgtgcaaaa cttcatcatc ggcggcggca atggcgggaa caatgcgttg    2940
gcgttccacg aagccctcga acggatccgc gaggattatc cggaggaagc gtggaaggtc    3000
atgctgaacc tcgtcgattc ccacgatacc gttcgctcgc tgacgaagct ggattatccg    3060
tcgtgggaag aagagaaact gaagatcgcg ccggaagcga gcgacaaggc gctgaagcag    3120
caggctttga tcgcgatatt ccaaatgggt tatccgggcg ccccgacgat ttattacggc    3180
gacgaagtcg gcgtaaccgg cacgaaggac ccggattcgc gcaggacgtt cccgtgggag    3240
cggatcgaaa gctccggcgg ctcctacaag gcggccggac gttatgcgga attgtatgag    3300
acgtacaaaa aggccgcagc cacgcgggcc gcaaacgacg tcttcgcgac cggggagctc    3360
aagctcgtct acgcgcaggg agacacgatc gcttacgccc ggaaaaatgc ggcgaaaggc    3420
gcgcttgtcg tcatcaacag cggcggcgag acggccgaga tgacggcgga tgtggccgga    3480
tttttgccgg acgggacgaa gctgagcgac cagcttggca gcggcatttc cggcaccgtc    3540
gagaacggcc gcattgccct gtccgtcccg gcgctgaccg gactcatgat gctgtccgac    3600
ggcgcgctga tttcggtcga ttcgccggac ggcttgaagg cgcaagccgg aaacggcagc    3660
gtggcgctta gctggaacgg ggtggcggat gcacaggcgt accgcgttta ccggtcgctg    3720
gtggagggcg gcgaagcgac gctgctcgga acggttacgg aaacgcagtt cacggatact    3780
tcggtcacca acggcgtgaa gtattattat ggggtaacgg cggtccgtac gacggcggac    3840
ggggccggcg aaagctcgct gcaggagatg gtcgcggcga ccctcgcatt tccggtaacg    3900
gccgtcggga acccgaccca ggccgaaaac atgacgatcg gcgtcggcaa ctcggccgac    3960
atcggcgtgg cggtcacggt tgcgggcttg acggacgacc cttcgcttgc cgggcgggaa    4020
gcgccgcgat tgagcgtgaa gctgatcgtg gccaagtcgg acgcttcgga acctgttgcc    4080
gagctttcga tggcgtacaa ggcggacgaa ggagacgcca agcggtacgc ggccaagttc    4140
gagccgaccg aagcgggcgt atacaactat ttcgtccggg tgtcgacgga taacgaagag    4200
acgtttacgg attcggcgac ggcttcgttc gaggcttatg cggacccgaa cgatacgatg    4260
ggacccgagg cgcctatctt aagcgaaatc acggtcgaat ccggtcaggc ggcgttgaac    4320
tggacgccgg cgggagaaga gacgtcgggc tacgaggtgt accggaaggc ggcgggggaa    4380
gcgacgttcc gcaagctggc cacgctgcgg acggccactt cctatatcga ttacacggtg    4440
acgaacggca cttcctacac gtacaaaatc gcggccttcg acgcggctta caaccggggc    4500
tggtcgaacg agcagaccgt tacgccgcaa ctggtgatgg tcgacgtgac gctgcggctg    4560
catttgccgg attacacgcc ggcgacggac tcgatttata ttgcgggaac cttgaacaac    4620
tggaacgcga gcggcggcaa gctgaccgtg ccgagcggcg cgaccggccg cagcgtcgtc    4680
caatatgcct tcaagatgat ggccggcaag acgatcgaat acaagtacac gcgcggcaca    4740
tgggcgacgg aagcactgac cagccatgcg cgggtcccga acgatacgac ggatgtgagc    4800
aactacgcgt acagcgacga aaaaaccaat atgaagctga cgatccgcaa tcagggcggc    4860
ggcaaaatga tcgtcgacga ctacgtgctg cgctgggcgg atatgccgat gatcgtcacg    4920
atgccgcgga tttcgtacgg cgaggatatc gcgttcgcca cgtccgacag cgccttcgat    4980
ctgaaggcga acgttccgta cggcgtcgcg ttcacgatca acggccagcc gctgccgaaa    5040
ggcgccatgg acgaccgggg caacgtctac gtggaaaaaa ttccgctgaa gccgggagag    5100
aatacgttca tcctccatat cgaaccgacg gcggagacgc tggcgcagga atggtataag    5160
gatcaaggcc gcaaaggcca ggcgaccaag tcgattgcgc tacgcattac gagaaccgcc    5220
ggagacggcg gcggtgggga cgacggagat ggcggtaacg gtggtagcgg cggagacggc    5280
ggcaacggcg gagatggcgg taatggcggt gacggcggca atgggggcga tggtcccggc    5340
actccggttg tcatcgtgcc tttgccctcc gctccttccg gatcgcttgt ggcccgacgg    5400
gacaccgaaa cggatgggca gggcaacacg gttaccgtcg cgacagtgga cggcgagaag    5460
ctggcggcgg ccatcgccgc tttgaaggac ggcaatggcg tggtcgagat ccggcttccg    5520
gatacgaagg gcagcgtccg gttcgcgatt ccgctgaaag cggcggctgc ggggaaagcc    5580
gcgtctccgg ccgcagccgt tcgcttggcg accgcagacg gcagctatgc gttgccgctg    5640
tccgtcctcg agccggagga gcttgcgaaa aagctgggcg tttccgtcga agacgtcacg    5700
cttacggtga gcataagcgt gctaacagga gcggaagccg agcgggcgca agcgctcgtc    5760
```

```
gaggcggcgg gcgcgaaagt gctgggcggt ccgatcgcgt atacggtcgt cgcctcggcc    5820
ggcgaacgcg aaacggagct ggcgagcttc ggttccgttt acgcggagcg gagcatgacg    5880
ctgaagggcg tcgtcgatcc gcggcatgcc tcggtgttcc gggttgacga cgaggcgaac    5940
aagctgagct tcgtgcccgc tttgtttacg gcggccggca acgggcaaac cgccgtcgcg    6000
ttcaagcggc cgggcaacag cgtctacatc gtcgcttcct cggaacgcga ctttgcggat    6060
gtcgcggatc attgggcgcg gccggatatc ctgctccttg ccaacaaggg catcgtccaa    6120
ggtgtccagg acggcttgtt cgctcccgac cggcgcatca cgcgcgcgga gttcgcgtcg    6180
ctgcttgtcc gctcgctcgg catggccgaa gcggacggaa ccggcggatt ccgcgacgtc    6240
cggccggacg attggtttgc cggcgtcgtg gcggcagcgg cgaaagcggg gcttgtgaac    6300
ggaacggaag acggctcgtt ccggcccggc gacggcatca cgcgggagca gatggccgtc    6360
atggccgcga aggcgctcgc cttcgccggc aaaccggctg cgggcggagc gccgcaggcg    6420
ggcgaaacgc ttgccgcgtt cgccgacggc gcggccgtca gcgcctgggc gcgggatgcc    6480
gtcgcccgga ccgtgcaggc cggcctcatg cagggcggca gcggcggccg gctcgcgccg    6540
tccgaaggcg cgacgcgagc gcaggcggtt gtgatgctga agcgcctgct ggagacgttg    6600
gcgtttatcg acaagtga                                                  6618



<210> 584
<211> 2205
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<221> DOMAIN
<222> (546)...(1109)
<223> Alpha amylase, catalytic domain

<221> DOMAIN
<222> (2018)...(2059)
<223> S-layer homology domain

<221> DOMAIN
<222> (2077)...(2119)
<223> S-layer homology domain

<221> DOMAIN
<222> (2147)...(2190)
<223> S-layer homology domain

<221> DOMAIN
<222> (1125)...(1203)
<223> Alpha amylase, C-terminal all-beta domain

<221> DOMAIN
<222> (1421)...(1502)
<223> Fibronectin type III domain

<221> DOMAIN
<222> (1206)...(1286)
<223> Fibronectin type III domain

<400> 584
Met Met Ala Val Ser Lys Lys Glu Ile Thr Gln Asp Met Phe Gln Pro
 1               5                  10                  15
Ala Val Ile Asp Lys Ser Leu Gln Glu Gln Val Ser Gly Pro Ser Ile
            20                  25                  30
Gly Phe Trp Ser Asp Gly Phe Arg Arg Leu Arg Lys Asn Lys Ala Ala
        35                  40                  45
Met Val Ser Ala Phe Ile Ile Leu Leu Ile Val Ile Met Ala Leu Val
    50                  55                  60
Gly Pro His Met Ser Lys Tyr Gly Ile Asp Asp Gln Asp Leu Lys Arg
65                  70                  75                  80
Ser Asn Leu Pro Pro Arg Ile Ser Ala Leu Ala Asn Ile His Trp Leu
                85                  90                  95
Pro Phe Asp Gly Lys Asn Glu Gln Gly Val Asp Val Tyr Lys Glu Arg
```

```
            100                 105                 110
Gly Ile Lys Gln Asp Leu Trp Phe Gly Ala Asp Lys Leu Gly Arg Asp
        115                 120                 125
Leu Trp Thr Arg Val Trp Lys Gly Thr Gln Ile Ser Leu Phe Ile Ala
    130                 135                 140
Val Val Ala Ala Ala Ala Asp Leu Ile Ile Gly Val Ala Tyr Gly Gly
145                 150                 155                 160
Ile Ser Ala Tyr Phe Gly Gly Gln Val Asp Asn Ile Met Gln Arg Ile
                165                 170                 175
Leu Glu Val Leu Asn Gly Ile Pro Asn Leu Val Ile Ile Ile Leu Leu
            180                 185                 190
Ile Ile Val Met Gln Pro Gly Ile Pro Ala Ile Ile Leu Ala Ile Ala
        195                 200                 205
Met Thr Gly Trp Ile Gly Met Ala Arg Ile Ile Arg Gly Glu Val Leu
    210                 215                 220
Lys Leu Lys Asn Leu Glu Phe Lys Leu Thr Lys Asn Gly Ser Tyr Ala
225                 230                 235                 240
Ser Gly Ile Pro Gly Asp Gly Glu Asn Leu Lys Leu Thr Leu Ser Gln
                245                 250                 255
Ala Ala Lys Val Asn Phe Tyr Val Asn Gly Asp Thr Gln Glu Ala Arg
            260                 265                 270
Ile Ser Pro Leu Pro Glu Gly Ser Thr Ala Ser Gly Ile Gln Ile Tyr
        275                 280                 285
Asn Pro Ile Leu Ser Ala Asp Lys Trp Pro Arg Leu Val Gly Ser Leu
    290                 295                 300
Gln Pro Leu Phe Gly Glu Pro Glu Trp Ser Pro Gly Glu Ala Lys Gln
305                 310                 315                 320
Phe Phe Val Asp Tyr Asn Phe Asn Asn Thr Val Tyr Lys Leu Gln Arg
                325                 330                 335
Ser Leu Pro Ala Gly Ser Tyr Glu Ala Lys Val Val Leu Gly Ala Asp
            340                 345                 350
Trp Ala Asp Gly Asn Asn Phe Gly Asn Gly Asn Ala Asn Ala Ala Phe
        355                 360                 365
Arg Leu Pro Asp Ala Ala Asp Val Thr Phe Ser Phe Asp Asn Ser Ala
    370                 375                 380
Ala Glu Arg Lys Leu Thr Val Asp Tyr Ala Val Gly Asp Gly Lys Tyr
385                 390                 395                 400
Asp Gly Lys Ile Asp Ala Gly Ala Ile Ala Phe Asp Ser Arg Gln Ala
                405                 410                 415
Thr Phe Lys Lys Pro Phe Gly Ala Ile Lys Gln Gly Glu Gln Asp Leu
            420                 425                 430
Thr Leu Arg Ile Ala Val Lys Arg Asp Asp Val Gln Met Ala Arg Val
        435                 440                 445
Glu Leu Thr Asp Ala His Gly Ala Ala Thr Asn Tyr Val Met Arg Lys
    450                 455                 460
Thr Thr Thr Val Gly Asp Arg Asp Tyr Tyr Glu Ala Val Ile Pro Lys
465                 470                 475                 480
Ser Val Phe Ala Gly Ile Gly Ile Trp Gly Tyr Lys Phe Val Leu Ile
                485                 490                 495
Asp Gly Ala Ala Lys Val Glu Tyr Gly Asp Asp Ser Leu Ser Gly Gly
            500                 505                 510
Ser Gly Glu Ala Val Ala Asp Gly Ala Val Pro Phe Asn Leu Thr Val
        515                 520                 525
Tyr Ala Ala Asp Tyr Lys Thr Pro Asp Trp Met Lys Thr Ala Val Val
    530                 535                 540
Tyr Gln Ile Phe Pro Asp Arg Phe Phe Asp Gly Asn Pro Asp Asn Asn
545                 550                 555                 560
Arg Ala Lys Leu Ala Asp Gly Tyr Arg Gly Val Arg Ser Glu Thr Gly
                565                 570                 575
Glu Gln Ala Gly Arg Ile Gly Lys Gln Glu Leu Gln Tyr Phe Asp Gly
            580                 585                 590
Gly Val Ala Asn Asp Pro Ala Pro Ser Gln Val Trp Gly Asp Trp Asn
        595                 600                 605
Asp Pro Pro Glu Asn Pro Asp Arg Thr Lys Pro Glu Asn Lys Pro Tyr
    610                 615                 620
Tyr Pro Asp Ala Lys Ser Asp Gly Ile Trp Thr Asn Glu Phe Tyr Gly
625                 630                 635                 640
Gly Asp Ile Gln Gly Ile Glu Gln Lys Leu Asp Tyr Leu Lys Gln Leu
```

645 650 655
Gly Val Thr Ala Ile Tyr Leu Asn Pro Val Ala Trp Ala Ala Ser Asn
660 665 670
His Lys Tyr Asp Ala Thr Asp Tyr Lys His Leu Asp Pro Met Phe Gly
675 680 685
Glu Pro Val Tyr Asn Thr Pro Gly Asp Pro Ser Ser Gly Leu Asn Tyr
690 695 700
Glu Glu Thr Arg Lys Lys Ser Asp Glu Val Tyr Arg Lys Phe Ala Lys
705 710 715 720
Ala Ala Ala Glu Ala Gly Ile Arg Ile Ile Asn Asp Gly Val Phe Asn
725 730 735
His Val Gly Asp Asp Ser Ile Tyr Phe Asp Arg Tyr Glu Lys Tyr Pro
740 745 750
Glu Ile Gly Ala Tyr Glu Tyr Trp Ser Lys Val Trp Lys Thr Lys Ala
755 760 765
Glu His Pro Gly Met Thr Arg Glu Glu Ala Glu Gln Arg Val Ile Ala
770 775 780
Ser Tyr Thr Ser Gln Ile Asn Pro Ala Thr Gly Arg His Tyr Gln Tyr
785 790 795 800
Pro Asp Asp Phe Gly Phe Thr Thr Trp Phe Thr Ile Gly Glu Asn Met
805 810 815
Ala Asp Asn Thr Arg Tyr Ser Tyr Asp Ala Trp Trp Gly Phe Asp Ser
820 825 830
Leu Pro Ala Met Asp Ala Lys Glu Pro Gln Pro Gly Asp Lys Asp Ala
835 840 845
Leu Glu Gly Val His Glu Trp Asn Asn Ile Asp Tyr Arg Asn His Val
850 855 860
Ile Gly Tyr Asp Leu Ser Gly Lys Ser Glu Ala Glu Ala Ser Ala Ala
865 870 875 880
Met Gln Asn Thr Ala Ser Gln Arg Trp Leu Trp Met Gly Ser Ser Gly
885 890 895
Trp Arg Leu Asp Val Ala Pro Asp Val Ser Ala Gly Thr Trp Arg Lys
900 905 910
Phe Arg Glu Ala Val Lys Ser Ala Ala Gly Arg Thr Asp Ala Asn Gly
915 920 925
Lys Pro Ile Asp Asp Pro Val Ile Leu Gly Glu Glu Trp Gly Val Ala
930 935 940
Thr His Tyr Leu Leu Gly Asp Gln Phe Asp Ser Ile Met Asn Tyr Arg
945 950 955 960
Phe Arg Glu Ala Val Gln Asn Phe Ile Ile Gly Gly Gly Asn Gly Gly
965 970 975
Asn Asn Ala Leu Ala Phe His Glu Ala Leu Glu Arg Ile Arg Glu Asp
980 985 990
Tyr Pro Glu Glu Ala Trp Lys Val Met Leu Asn Leu Val Asp Ser His
995 1000 1005
Asp Thr Val Arg Ser Leu Thr Lys Leu Asp Tyr Pro Ser Trp Glu Glu
1010 1015 1020
Glu Lys Leu Lys Ile Ala Pro Glu Ala Ser Asp Lys Ala Leu Lys Gln
1025 1030 1035 1040
Gln Ala Leu Ile Ala Ile Phe Gln Met Gly Tyr Pro Gly Ala Pro Thr
1045 1050 1055
Ile Tyr Tyr Gly Asp Glu Val Gly Val Thr Gly Thr Lys Asp Pro Asp
1060 1065 1070
Ser Arg Arg Thr Phe Pro Trp Glu Arg Ile Glu Ser Ser Gly Gly Ser
1075 1080 1085
Tyr Lys Ala Ala Gly Arg Tyr Ala Glu Leu Tyr Glu Thr Tyr Lys Lys
1090 1095 1100
Ala Ala Ala Thr Arg Ala Ala Asn Asp Val Phe Ala Thr Gly Glu Leu
1105 1110 1115 1120
Lys Leu Val Tyr Ala Gln Gly Asp Thr Ile Ala Tyr Ala Arg Lys Asn
1125 1130 1135
Ala Ala Lys Gly Ala Leu Val Val Ile Asn Ser Gly Gly Glu Thr Ala
1140 1145 1150
Glu Met Thr Ala Asp Val Ala Gly Phe Leu Pro Asp Gly Thr Lys Leu
1155 1160 1165
Ser Asp Gln Leu Gly Ser Gly Ile Ser Gly Thr Val Glu Asn Gly Arg
1170 1175 1180
Ile Ala Leu Ser Val Pro Ala Leu Thr Gly Leu Met Met Leu Ser Asp

```
1185                1190                1195                1200
Gly Ala Leu Ile Ser Val Asp Ser Pro Asp Gly Leu Lys Ala Gln Ala
            1205                1210                1215
Gly Asn Gly Ser Val Ala Leu Ser Trp Asn Gly Val Ala Asp Ala Gln
        1220                1225                1230
Ala Tyr Arg Val Tyr Arg Ser Leu Val Glu Gly Gly Glu Ala Thr Leu
    1235                1240                1245
Leu Gly Thr Val Thr Glu Thr Gln Phe Thr Asp Thr Ser Val Thr Asn
1250                1255                1260
Gly Val Lys Tyr Tyr Tyr Gly Val Thr Ala Val Arg Thr Thr Ala Asp
1265                1270                1275                1280
Gly Ala Gly Glu Ser Ser Leu Gln Glu Met Val Ala Ala Thr Leu Ala
            1285                1290                1295
Phe Pro Val Thr Ala Val Gly Asn Pro Thr Gln Ala Glu Asn Met Thr
        1300                1305                1310
Ile Gly Val Gly Asn Ser Ala Asp Ile Gly Val Ala Val Thr Val Ala
    1315                1320                1325
Gly Leu Thr Asp Asp Pro Ser Leu Ala Gly Arg Glu Ala Pro Arg Leu
1330                1335                1340
Ser Val Lys Leu Ile Val Ala Lys Ser Asp Ala Ser Glu Pro Val Ala
1345                1350                1355                1360
Glu Leu Ser Met Ala Tyr Lys Ala Asp Glu Gly Asp Ala Lys Arg Tyr
            1365                1370                1375
Ala Ala Lys Phe Glu Pro Thr Glu Ala Gly Val Tyr Asn Tyr Phe Val
        1380                1385                1390
Arg Val Ser Thr Asp Asn Glu Glu Thr Phe Thr Asp Ser Ala Thr Ala
    1395                1400                1405
Ser Phe Glu Ala Tyr Ala Asp Pro Asn Asp Thr Met Gly Pro Glu Ala
1410                1415                1420
Pro Ile Leu Ser Glu Ile Thr Val Glu Ser Gly Gln Ala Ala Leu Asn
1425                1430                1435                1440
Trp Thr Pro Ala Gly Glu Glu Thr Ser Gly Tyr Glu Val Tyr Arg Lys
            1445                1450                1455
Ala Ala Gly Glu Ala Thr Phe Arg Lys Leu Ala Thr Leu Arg Thr Ala
        1460                1465                1470
Thr Ser Tyr Ile Asp Tyr Thr Val Thr Asn Gly Thr Ser Tyr Thr Tyr
    1475                1480                1485
Lys Ile Ala Ala Phe Asp Ala Ala Tyr Asn Arg Gly Trp Ser Asn Glu
1490                1495                1500
Gln Thr Val Thr Pro Gln Leu Val Met Val Asp Val Thr Leu Arg Leu
1505                1510                1515                1520
His Leu Pro Asp Tyr Thr Pro Ala Thr Asp Ser Ile Tyr Ile Ala Gly
            1525                1530                1535
Thr Leu Asn Asn Trp Asn Ala Ser Gly Gly Lys Leu Thr Val Pro Ser
        1540                1545                1550
Gly Ala Thr Gly Arg Ser Val Val Gln Tyr Ala Phe Lys Met Met Ala
    1555                1560                1565
Gly Lys Thr Ile Glu Tyr Lys Tyr Thr Arg Gly Thr Trp Ala Thr Glu
1570                1575                1580
Ala Leu Thr Ser His Ala Arg Val Pro Asn Asp Thr Thr Asp Val Ser
1585                1590                1595                1600
Asn Tyr Ala Tyr Ser Asp Glu Lys Thr Asn Met Lys Leu Thr Ile Arg
            1605                1610                1615
Asn Gln Gly Gly Gly Lys Met Ile Val Asp Asp Tyr Val Leu Arg Trp
        1620                1625                1630
Ala Asp Met Pro Met Ile Val Thr Met Pro Arg Ile Ser Tyr Gly Glu
    1635                1640                1645
Asp Ile Ala Phe Ala Thr Ser Asp Ser Ala Phe Asp Leu Lys Ala Asn
1650                1655                1660
Val Pro Tyr Gly Val Ala Phe Thr Ile Asn Gly Gln Pro Leu Pro Lys
1665                1670                1675                1680
Gly Ala Met Asp Asp Arg Gly Asn Val Tyr Val Glu Lys Ile Pro Leu
            1685                1690                1695
Lys Pro Gly Glu Asn Thr Phe Ile Leu His Ile Glu Pro Thr Ala Glu
        1700                1705                1710
Thr Leu Ala Gln Glu Trp Tyr Lys Asp Gln Gly Arg Lys Gly Gln Ala
    1715                1720                1725
Thr Lys Ser Ile Ala Leu Arg Ile Thr Arg Thr Ala Gly Asp Gly Gly
```

1730 1735 1740
Gly Gly Asp Asp Gly Asp Gly Gly Asn Gly Gly Ser Gly Gly Asp Gly
1745 1750 1755 1760
Gly Asn Gly Gly Asp Gly Gly Asn Gly Gly Asp Gly Gly Asn Gly Gly
1765 1770 1775
Asp Gly Pro Gly Thr Pro Val Val Ile Val Pro Leu Pro Ser Ala Pro
1780 1785 1790
Ser Gly Ser Leu Val Ala Arg Arg Asp Thr Glu Thr Asp Gly Gln Gly
1795 1800 1805
Asn Thr Val Thr Val Ala Thr Val Asp Gly Glu Lys Leu Ala Ala Ala
1810 1815 1820
Ile Ala Ala Leu Lys Asp Gly Asn Gly Val Val Glu Ile Arg Leu Pro
1825 1830 1835 1840
Asp Thr Lys Gly Ser Val Arg Phe Ala Ile Pro Leu Lys Ala Ala Ala
1845 1850 1855
Ala Gly Lys Ala Ala Ser Pro Ala Ala Ala Val Arg Leu Ala Thr Ala
1860 1865 1870
Asp Gly Ser Tyr Ala Leu Pro Leu Ser Val Leu Glu Pro Glu Glu Leu
1875 1880 1885
Ala Lys Lys Leu Gly Val Ser Val Glu Asp Val Thr Leu Thr Val Ser
1890 1895 1900
Ile Ser Val Leu Thr Gly Ala Glu Ala Glu Arg Ala Gln Ala Leu Val
1905 1910 1915 1920
Glu Ala Ala Gly Ala Lys Val Leu Gly Gly Pro Ile Ala Tyr Thr Val
1925 1930 1935
Val Ala Ser Ala Gly Glu Arg Glu Thr Glu Leu Ala Ser Phe Gly Ser
1940 1945 1950
Val Tyr Ala Glu Arg Ser Met Thr Leu Lys Gly Val Val Asp Pro Arg
1955 1960 1965
His Ala Ser Val Phe Arg Val Asp Asp Glu Ala Asn Lys Leu Ser Phe
1970 1975 1980
Val Pro Ala Leu Phe Thr Ala Ala Gly Asn Gly Gln Thr Ala Val Ala
1985 1990 1995 2000
Phe Lys Arg Pro Gly Asn Ser Val Tyr Ile Val Ala Ser Ser Glu Arg
2005 2010 2015
Asp Phe Ala Asp Val Ala Asp His Trp Ala Arg Pro Asp Ile Leu Leu
2020 2025 2030
Leu Ala Asn Lys Gly Ile Val Gln Gly Val Gln Asp Gly Leu Phe Ala
2035 2040 2045
Pro Asp Arg Arg Ile Thr Arg Ala Glu Phe Ala Ser Leu Leu Val Arg
2050 2055 2060
Ser Leu Gly Met Ala Glu Ala Asp Gly Thr Gly Gly Phe Arg Asp Val
2065 2070 2075 2080
Arg Pro Asp Asp Trp Phe Ala Gly Val Val Ala Ala Ala Ala Lys Ala
2085 2090 2095
Gly Leu Val Asn Gly Thr Glu Asp Gly Ser Phe Arg Pro Gly Asp Gly
2100 2105 2110
Ile Thr Arg Glu Gln Met Ala Val Met Ala Ala Lys Ala Leu Ala Phe
2115 2120 2125
Ala Gly Lys Pro Ala Ala Gly Gly Ala Pro Gln Ala Gly Glu Thr Leu
2130 2135 2140
Ala Ala Phe Ala Asp Gly Ala Ala Val Ser Ala Trp Ala Arg Asp Ala
2145 2150 2155 2160
Val Ala Arg Thr Val Gln Ala Gly Leu Met Gln Gly Gly Ser Gly Gly
2165 2170 2175
Arg Leu Ala Pro Ser Glu Gly Ala Thr Arg Ala Gln Ala Val Val Met
2180 2185 2190
Leu Lys Arg Leu Leu Glu Thr Leu Ala Phe Ile Asp Lys
2195 2200 2205

<210> 585
<211> 1860
<212> DNA
<213> Bacteria

<400> 585
atgtccctat tcaaaaaaag ctttccgtgg attttatccc tacttctttt gtttttgttt 60
attgctcctt tttccattca aacagaaaaa gtccgagctg gaagtgtgcc ggtaaacggc 120

```
acaatgatgc aatatttcga atggtacctt ccagacgatg gaacactatg gacgaaagta     180
gcaaataacg ctcaatcttt agcgaatctt ggcattactg ccctttggct tccccctgcc     240
tataaaggaa caagcagcag tgacgttgga tatggcgttt atgatttata tgaccttgga     300
gagtttaatc aaaaaggaac tgtccgaaca aaatacggga caaaaacaca atatatccaa     360
gcaatccaag cggcccatac agcagggatg caagtatatg cagatgtcgt ctttaaccat     420
aaagccggtg cagatggaac agaactagtt gatgcagtag aagtaaatcc ttctgaccgc     480
aatcaagaaa tatcaggaac atatcaaatc caagcgtgga caaaatttga ttttcctggt     540
cgtggaaaca cctattctag ttttaaatgg cgttggtatc atttcgatgg aacggactgg     600
gatgagagta gaaaactaaa tcgtatttac aagttccgcg gcacgggaaa agcatgggat     660
tgggaagtag atacagaaaa cgggaattat gactatctca tgtatgcaga tttggatatg     720
gatcatccag aggttgtatc cgaactaaaa aattggggaa agtggtatgt aaccacaacc     780
aatatcgacg gattccgtct ggatgcagtg aagcatatta aatatagctt tttcccggac     840
tggctatcgt acgtacgaac ccaaacacaa aagcctcttt tttccgttgg ggaattttgg     900
agctatgaca ttagcaagct gcacaactat attacaaaga cgaacggctc tatgtcccta     960
ttcgatgccc cgctgcataa caatttttat atagcatcga aatcaggcgg ttattttgat    1020
atgcgcacat tactcaacaa cacattgatg aaagatcagc ctacattagc agtcacatta    1080
gtggataatc acgatactga gccagggcaa tctttgcagt catgggtcga gccatggttt    1140
aaaccgttag cttacgcatt tatcttgacc cgccaagaag gttatccttg tgtcttttat    1200
ggagattact atggtattcc aaaatacaac attcctgcgc tgaaaagcaa acttgatccg    1260
ctgttaattg ccagaagaga ttatgcctat ggaacacagc acgactatat tgacagtgcg    1320
gatattatcg gttggacgcg ggaaggagtg gctgaaaaag caaattcagg actggctgca    1380
ctcattaccg acgggcctgg cggaagcaaa tggatgtatg ttggaaaaca acacgctggc    1440
aaaacgtttt atgatttaac cggcaatcga agtgatacag tgacaatcaa tgctgatgga    1500
tggggagaat ttaaagtcaa tggagggtct gtatccatat gggttccaaa aacatcaacc    1560
acttcccaaa taacatttac tgtaaataat gccacaaccg tttggggaca aaatgtatac    1620
gttgtcggga atatttcgca gctggggaac tgggatccag tcaacgcagt tcaaatgacg    1680
ccgtcttctt atccaacatg ggtagtgaca gtccctcttc cacaaagtca aaacatacaa    1740
tttaaattta tcaaaaaaga tggatcagga aatgtcattt gggagaatat atcgaatcga    1800
acatacaccg tcccaactgc tgcatccgga gcatatacag ccagctggaa cgtgccctag    1860

<210> 586
<211> 619
<212> PRT
<213> Bacteria

<220>
<221> SIGNAL
<222> (1)...(28)

<221> DOMAIN
<222> (521)...(614)
<223> Starch binding domain

<221> DOMAIN
<222> (40)...(425)
<223> Alpha amylase, catalytic domain

<400> 586
Met Ser Leu Phe Lys Lys Ser Phe Pro Trp Ile Leu Ser Leu Leu Leu
 1               5                  10                  15
Leu Phe Leu Phe Ile Ala Pro Phe Ser Ile Gln Thr Glu Lys Val Arg
            20                  25                  30
Ala Gly Ser Val Pro Val Asn Gly Thr Met Met Gln Tyr Phe Glu Trp
        35                  40                  45
Tyr Leu Pro Asp Asp Gly Thr Leu Trp Thr Lys Val Ala Asn Asn Ala
    50                  55                  60
Gln Ser Leu Ala Asn Leu Gly Ile Thr Ala Leu Trp Leu Pro Pro Ala
65                  70                  75                  80
Tyr Lys Gly Thr Ser Ser Ser Asp Val Gly Tyr Gly Val Tyr Asp Leu
                85                  90                  95
Tyr Asp Leu Gly Glu Phe Asn Gln Lys Gly Thr Val Arg Thr Lys Tyr
            100                 105                 110
Gly Thr Lys Thr Gln Tyr Ile Gln Ala Ile Gln Ala Ala His Thr Ala
        115                 120                 125
Gly Met Gln Val Tyr Ala Asp Val Val Phe Asn His Lys Ala Gly Ala
    130                 135                 140
Asp Gly Thr Glu Leu Val Asp Ala Val Glu Val Asn Pro Ser Asp Arg
145                 150                 155                 160
```

```
Asn Gln Glu Ile Ser Gly Thr Tyr Gln Ile Gln Ala Trp Thr Lys Phe
                165                 170                 175
Asp Phe Pro Gly Arg Gly Asn Thr Tyr Ser Ser Phe Lys Trp Arg Trp
            180                 185                 190
Tyr His Phe Asp Gly Thr Asp Trp Asp Glu Ser Arg Lys Leu Asn Arg
        195                 200                 205
Ile Tyr Lys Phe Arg Gly Thr Gly Lys Ala Trp Asp Trp Glu Val Asp
    210                 215                 220
Thr Glu Asn Gly Asn Tyr Asp Tyr Leu Met Tyr Ala Asp Leu Asp Met
225                 230                 235                 240
Asp His Pro Glu Val Val Ser Glu Leu Lys Asn Trp Gly Lys Trp Tyr
                245                 250                 255
Val Thr Thr Thr Asn Ile Asp Gly Phe Arg Leu Asp Ala Val Lys His
            260                 265                 270
Ile Lys Tyr Ser Phe Phe Pro Asp Trp Leu Ser Tyr Val Arg Thr Gln
        275                 280                 285
Thr Gln Lys Pro Leu Phe Ser Val Gly Glu Phe Trp Ser Tyr Asp Ile
    290                 295                 300
Ser Lys Leu His Asn Tyr Ile Thr Lys Thr Asn Gly Ser Met Ser Leu
305                 310                 315                 320
Phe Asp Ala Pro Leu His Asn Asn Phe Tyr Ile Ala Ser Lys Ser Gly
                325                 330                 335
Gly Tyr Phe Asp Met Arg Thr Leu Leu Asn Asn Thr Leu Met Lys Asp
            340                 345                 350
Gln Pro Thr Leu Ala Val Thr Leu Val Asp Asn His Asp Thr Glu Pro
        355                 360                 365
Gly Gln Ser Leu Gln Ser Trp Val Glu Pro Trp Phe Lys Pro Leu Ala
    370                 375                 380
Tyr Ala Phe Ile Leu Thr Arg Gln Glu Gly Tyr Pro Cys Val Phe Tyr
385                 390                 395                 400
Gly Asp Tyr Tyr Gly Ile Pro Lys Tyr Asn Ile Pro Ala Leu Lys Ser
                405                 410                 415
Lys Leu Asp Pro Leu Leu Ile Ala Arg Arg Asp Tyr Ala Tyr Gly Thr
            420                 425                 430
Gln His Asp Tyr Ile Asp Ser Ala Asp Ile Ile Gly Trp Thr Arg Glu
        435                 440                 445
Gly Val Ala Glu Lys Ala Asn Ser Gly Leu Ala Ala Leu Ile Thr Asp
    450                 455                 460
Gly Pro Gly Gly Ser Lys Trp Met Tyr Val Gly Lys Gln His Ala Gly
465                 470                 475                 480
Lys Thr Phe Tyr Asp Leu Thr Gly Asn Arg Ser Asp Thr Val Thr Ile
                485                 490                 495
Asn Ala Asp Gly Trp Gly Glu Phe Lys Val Asn Gly Gly Ser Val Ser
            500                 505                 510
Ile Trp Val Pro Lys Thr Ser Thr Thr Ser Gln Ile Thr Phe Thr Val
        515                 520                 525
Asn Asn Ala Thr Thr Val Trp Gly Gln Asn Val Tyr Val Val Gly Asn
    530                 535                 540
Ile Ser Gln Leu Gly Asn Trp Asp Pro Val Asn Ala Val Gln Met Thr
545                 550                 555                 560
Pro Ser Ser Tyr Pro Thr Trp Val Val Thr Val Pro Leu Pro Gln Ser
                565                 570                 575
Gln Asn Ile Gln Phe Lys Phe Ile Lys Lys Asp Gly Ser Gly Asn Val
            580                 585                 590
Ile Trp Glu Asn Ile Ser Asn Arg Thr Tyr Thr Val Pro Thr Ala Ala
        595                 600                 605
Ser Gly Ala Tyr Thr Ala Ser Trp Asn Val Pro
    610                 615

<210> 587
<211> 4111
<212> DNA
<213> Thermomyces lanuginosus ATCC 200065

<220>
<221> CDS
<222> (702)...(942)
<223> Exon
```

<221> CDS
<222> (1007)...(1290)
<223> Exon

<221> CDS
<222> (1352)...(1447)
<223> Exon

<221> CDS
<222> (1528)...(2166)
<223> Exon

<221> CDS
<222> (2224)...(2817)
<223> Exon

<400> 587

```
tccaaagaat tgaacgtcgt cggatgatcg tgatggtcag ttttgaggta tcttccgaag     60
agccaatttg gattacccgg cagtccggca gtccgacggt ccagacaagg atcaatcagg    120
gcctggcact attttagcat gtcggttagt ccccaatcgg ctctgcattt gacagcggac    180
atctgccgat caggaacgaa acggtgcctg ccgttgttgc cctgtgagct ctttatgcat    240
gtcaattcat aataatcccg ataataggcc gcagctattc caccctgggg aaagggtttg    300
ttgaatgtcg gcgatttggg gtacatctgg gggttcggca cttcagtgtc tccaggtccg    360
aatttcatgc attcatgcat gaaaatgcgg agggaacctc ggcactgtca ccatccccgt    420
caaactacgg cttttggcgg gctacattgg ctacaatggc cacaatggct acaacggcct    480
gtcaaggtat ccattaattt ctccgaagct cctgcaacgt ttcgacaccc taaatttccg    540
tttccgctac ggagagcgga gtaccggata ctttttggac ttttcatgtg cttctgtcgg    600
gcattcaaag accgggcctg tcctaacgtc caggggatcg tttggcagat cttccctcat    660
gcgttcacag agcgaagttt gtttaattga gtcatcaaaa c atg tta ttc caa ccg    716
                                              Met Leu Phe Gln Pro
                                               1               5

act ttg tgc gcg gcc ctt gga ctc gcc gcc ttg atc gtc caa ggc gga      764
Thr Leu Cys Ala Ala Leu Gly Leu Ala Ala Leu Ile Val Gln Gly Gly
                 10                  15                  20

gaa gcc aga cct gaa aca acc gtc cca cat gca acg ggc tcg ctc gac      812
Glu Ala Arg Pro Glu Thr Thr Val Pro His Ala Thr Gly Ser Leu Asp
             25                  30                  35

gac ttc ctc gcc gca cag agt ccg att gct ttc caa ggc atc ctg aac      860
Asp Phe Leu Ala Ala Gln Ser Pro Ile Ala Phe Gln Gly Ile Leu Asn
         40                  45                  50

aat atc ggg cct agc gga gcg tac tcg gaa ggt gtc aat ccg ggt gtg      908
Asn Ile Gly Pro Ser Gly Ala Tyr Ser Glu Gly Val Asn Pro Gly Val
     55                  60                  65

gtc att gcg agt cca agt aaa caa gat ccc gac t gtatgcctgc             952
Val Ile Ala Ser Pro Ser Lys Gln Asp Pro Asp
 70                  75                  80

tctggaaatt tttcaattct gttggcagga ctctctctct aatatggcac atag ac      1008
                                                          Tyr

ttt tac acc tgg gtg cgc gac gct gct ctc act gtc caa tat ctg gtg     1056
Phe Tyr Thr Trp Val Arg Asp Ala Ala Leu Thr Val Gln Tyr Leu Val
             85                  90                  95

gag gag ctg gtt gca gga aat gcc agt ctt cag ttc ctc att cag gac     1104
Glu Glu Leu Val Ala Gly Asn Ala Ser Leu Gln Phe Leu Ile Gln Asp
        100                 105                 110

tac atc agc tcc cag gca cga ctg cag acg gtg gaa aat cca tcc ggc     1152
Tyr Ile Ser Ser Gln Ala Arg Leu Gln Thr Val Glu Asn Pro Ser Gly
    115                 120                 125
```

```
tcc ctc tcg tcg ggt ggt cta gga gag ccc aag ttt cat gtc gac gag     1200
Ser Leu Ser Ser Gly Gly Leu Gly Glu Pro Lys Phe His Val Asp Glu
130                 135                 140                 145

acc gcc ttt acg gac tcc tgg ggc cga cca cag cgg gac ggc ccg cct     1248
Thr Ala Phe Thr Asp Ser Trp Gly Arg Pro Gln Arg Asp Gly Pro Pro
                150                 155                 160

ctc cgc gcc att gcc atg att tcg ttt gcc aat tac ctg att             1290
Leu Arg Ala Ile Ala Met Ile Ser Phe Ala Asn Tyr Leu Ile
            165                 170                 175

gtaagtcaga tttcccatca tgcgagtaaa ttgacatgga tgtgctcagt gtagttttca   1350
g gac aac ggt cat caa tcg act gtg gag gac atc atc tgg ccg att gtt   1399
  Asp Asn Gly His Gln Ser Thr Val Glu Asp Ile Ile Trp Pro Ile Val
                180                 185                 190

cgc aat gac ttg tcc tat gtc tcg cag cat tgg aac gaa aca act ttt     1447
Arg Asn Asp Leu Ser Tyr Val Ser Gln His Trp Asn Glu Thr Thr Phe
            195                 200                 205

ggtatgtgct tacgccgtac tagttgattg gagagtttgg attataggag agcctcaagc   1507
taatacggag tttttccgaa gac atc tgg gag gaa gtc cat agc tca tcg ttt   1560
                     Asp Ile Trp Glu Glu Val His Ser Ser Ser Phe
                         210                 215

ttc acc acg gct gtc cag tac cgt gct ctg gtc caa ggc agt gcc ttg     1608
Phe Thr Thr Ala Val Gln Tyr Arg Ala Leu Val Gln Gly Ser Ala Leu
    220                 225                 230

gct agc aag ctc ggc cat acc tgc gac aac tgc ggg tcc caa gca ccg     1656
Ala Ser Lys Leu Gly His Thr Cys Asp Asn Cys Gly Ser Gln Ala Pro
235                 240                 245                 250

cag atc ctt tgc ttc ctg cag tcg tat tgg acc ggg tcg cac atc tta     1704
Gln Ile Leu Cys Phe Leu Gln Ser Tyr Trp Thr Gly Ser His Ile Leu
                255                 260                 265

gcc aac acc ggt ggc ggc cgc tcg gga aag gac gtc agc acg atc ctc     1752
Ala Asn Thr Gly Gly Gly Arg Ser Gly Lys Asp Val Ser Thr Ile Leu
            270                 275                 280

ggc gtc att ggc tcg ttt gat ccg aac gcc gac tgt gat gac gtt acc     1800
Gly Val Ile Gly Ser Phe Asp Pro Asn Ala Asp Cys Asp Asp Val Thr
        285                 290                 295

ttc cag ccc tgc tcg gcc cgg gct ctt gca aat cac aag cag gtc gtt     1848
Phe Gln Pro Cys Ser Ala Arg Ala Leu Ala Asn His Lys Gln Val Val
    300                 305                 310

gac agc ttc cgc agt atc tat gcc atc aac gct ggc atc ccg tca ggg     1896
Asp Ser Phe Arg Ser Ile Tyr Ala Ile Asn Ala Gly Ile Pro Ser Gly
315                 320                 325                 330

tcg gct gtt gcg gtt gga cgt tat ccc gag gat gtc tat cag ggt gga     1944
Ser Ala Val Ala Val Gly Arg Tyr Pro Glu Asp Val Tyr Gln Gly Gly
                335                 340                 345

cac ccc tgg tac cta aca acg gct gcg gcg gcg gag cag ctt tac gac     1992
His Pro Trp Tyr Leu Thr Thr Ala Ala Ala Ala Glu Gln Leu Tyr Asp
            350                 355                 360

gcc att tac cag tgg aac cat gta ggg cac atc gac atc aat gct gtc     2040
Ala Ile Tyr Gln Trp Asn His Val Gly His Ile Asp Ile Asn Ala Val
        365                 370                 375

aat ctg gac ttc ttc aag agc att tat ccg tca gcc gcc gag ggc aca     2088
```

```
Asn Leu Asp Phe Phe Lys Ser Ile Tyr Pro Ser Ala Ala Glu Gly Thr
    380                 385                 390

tac aca tca gac tct tca aca ttt caa gac att ata tct gct gta cgg     2136
Tyr Thr Ser Asp Ser Ser Thr Phe Gln Asp Ile Ile Ser Ala Val Arg
395                 400                 405                 410

acc tat gcg gac ggg ttt ctc agc gta att gtaagtccaa accttcgaaa       2186
Thr Tyr Ala Asp Gly Phe Leu Ser Val Ile
                415                 420

acgaatgcct caagtcttcc actgacattt tgcgcag gag aaa tac act ccg ccg    2241
                                        Glu Lys Tyr Thr Pro Pro
                                                    425

gat aac ttg ctt gcc gag cag ttc cac cgg gag acg ggc att cca cta     2289
Asp Asn Leu Leu Ala Glu Gln Phe His Arg Glu Thr Gly Ile Pro Leu
            430                 435                 440

tcg gca gct tct ctg aca tgg tcc tac gcc gcg ctc aac acg gcc gcg     2337
Ser Ala Ala Ser Leu Thr Trp Ser Tyr Ala Ala Leu Asn Thr Ala Ala
        445                 450                 455

cag cgg cga gcg tca atc gtg ccc tca ccg tgg aac tct aac agc aca     2385
Gln Arg Arg Ala Ser Ile Val Pro Ser Pro Trp Asn Ser Asn Ser Thr
    460                 465                 470

gat ctc ccg gac aaa tgc tcg gca acc tcg gca aca ggg ccg tat gcc     2433
Asp Leu Pro Asp Lys Cys Ser Ala Thr Ser Ala Thr Gly Pro Tyr Ala
475                 480                 485                 490

acg ccc aca aac acg gca tgg cca acc act acg cag cca ccg gag cgg     2481
Thr Pro Thr Asn Thr Ala Trp Pro Thr Thr Thr Gln Pro Pro Glu Arg
                495                 500                 505

ccg gca tgc aca ccg ccg tcg gaa gta aca ctc acc ttc aac gcg ctc     2529
Pro Ala Cys Thr Pro Pro Ser Glu Val Thr Leu Thr Phe Asn Ala Leu
            510                 515                 520

gtc gac acc gcg ttt ggc cag aat att tat ctc gtg ggc tcc att ccg     2577
Val Asp Thr Ala Phe Gly Gln Asn Ile Tyr Leu Val Gly Ser Ile Pro
        525                 530                 535

gag ctc gga tcg tgg gat ccg gcc aac gcc ctc ttg atg agc gca aag     2625
Glu Leu Gly Ser Trp Asp Pro Ala Asn Ala Leu Leu Met Ser Ala Lys
    540                 545                 550

agc tgg act agc gga aat ccg gtc tgg acg cta tcc att tcc ctt cca     2673
Ser Trp Thr Ser Gly Asn Pro Val Trp Thr Leu Ser Ile Ser Leu Pro
555                 560                 565                 570

gca gga acc tct ttt gag tac aag ttc att cga aag gat gat ggt tcc     2721
Ala Gly Thr Ser Phe Glu Tyr Lys Phe Ile Arg Lys Asp Asp Gly Ser
                575                 580                 585

tcg gat gtt gtc tgg gaa agt gac ccg aat cgt tcg tac aac gtg ccg     2769
Ser Asp Val Val Trp Glu Ser Asp Pro Asn Arg Ser Tyr Asn Val Pro
            590                 595                 600

aag gat tgc ggt gcc aac acg gcc acc gtg aat tct tgg tgg cga tga     2817
Lys Asp Cys Gly Ala Asn Thr Ala Thr Val Asn Ser Trp Trp Arg  *
        605                 610                 615

accaacttgt ttctgtccac actccgccct gttgtcagtt cctggtcgta gatcgataaa   2877
atatgacttg tgtacttgac aaagaaatga tgtaaaagcg ttctgttatg tagtaggtta   2937
gcacttttcc ctagtaggga gtactccgta ggtatgccga taccgaactc cgaccggagt   2997
aaaactaacg tgtcgagtat cgcgatgttg cgcgtgggga gtaaggacga ttagggctga   3057
atgctgcaga tcctttcctt caccgtctca gagacacgga gtcaggtttg accatggccg   3117
```

```
cggtgtgaag ctgttgatgc tcattttcct gtccaatatt tcaaggacaa cgttgccaac   3177
atcaatccca gcaatgagtc ctcgtgttga ctgtgctgcc ctgggatcag agacttcggg   3237
aatacattga agttagcaaa acaatgccgt ccactaacta tgtgcgctca cgtctaccaa   3297
gtgcacgatc cgttctgtgg aaggggaagt tatcctggta caggcttgct tgcgattctg   3357
gatcgactgc gaaaaaaaga ggggacatgc gtctggaaag ttgcgattca aaacgatgtc   3417
atcagctcat tgcctcttga aatcctgctc gagatcagag attatcttga actaaacgat   3477
attctccgaa gccggatggt gcgtatcgtc ttgttgctcg atgttgaaag agacattttt   3537
atttgacagg agcaaacgtc cgaacaagtc tcgaaacaat agcgctctat tttctcaagc   3597
ggtccagtca tgagaccttt tctgcgggag gctcttgcgc tcctggacat gaaaggcacg   3657
gagatatccg ccacggacgt tatggcttat tttcgatggt acggtggatt gaaatatgga   3717
aagccagtaa agaaaatgtt tctcccatgg ccagaaggcc ggataagatt ggagcaaagg   3777
attgagattt attcgcggcg gctatattat catgaatgga agaagagagc agtcggaatg   3837
ttggatctgg aaactagaaa gaggcggctt tggcccgcaa acgataacac gtactttgaa   3897
ttctttgcat ctgatcgata tcttttgatc tcggggtaag ttgaggacag gatagtagat   3957
tcctgttggc tcgacaggtc gtctaactgc agtgctagga aacgtctgat tgcgtgggat   4017
atacaggagt cgcggaaagt tggggagatt agtggtgagc cctggcaggt gaacatcgtg   4077
caggataagg tcgtcatgtt tgaacgaaag tcaa                               4111
```

<210> 588
<211> 80
<212> PRT
<213> Thermomyces lanuginosus ATCC 200065

<400> 588

```
Met Leu Phe Gln Pro Thr Leu Cys Ala Ala Leu Gly Leu Ala Ala Leu
 1               5                  10                  15
Ile Val Gln Gly Gly Glu Ala Arg Pro Glu Thr Thr Val Pro His Ala
            20                  25                  30
Thr Gly Ser Leu Asp Asp Phe Leu Ala Ala Gln Ser Pro Ile Ala Phe
        35                  40                  45
Gln Gly Ile Leu Asn Asn Ile Gly Pro Ser Gly Ala Tyr Ser Glu Gly
    50                  55                  60
Val Asn Pro Gly Val Val Ile Ala Ser Pro Ser Lys Gln Asp Pro Asp
65                  70                  75                  80
```

<210> 589
<211> 95
<212> PRT
<213> Thermomyces lanuginosus ATCC 200065

<400> 589

```
Tyr Phe Tyr Thr Trp Val Arg Asp Ala Ala Leu Thr Val Gln Tyr Leu
 1               5                  10                  15
Val Glu Glu Leu Val Ala Gly Asn Ala Ser Leu Gln Phe Leu Ile Gln
            20                  25                  30
Asp Tyr Ile Ser Ser Gln Ala Arg Leu Gln Thr Val Glu Asn Pro Ser
        35                  40                  45
Gly Ser Leu Ser Ser Gly Gly Leu Gly Glu Pro Lys Phe His Val Asp
    50                  55                  60
Glu Thr Ala Phe Thr Asp Ser Trp Gly Arg Pro Gln Arg Asp Gly Pro
65                  70                  75                  80
Pro Leu Arg Ala Ile Ala Met Ile Ser Phe Ala Asn Tyr Leu Ile
                85                  90                  95
```

<210> 590
<211> 32
<212> PRT
<213> Thermomyces lanuginosus ATCC 200065

<400> 590

```
Asp Asn Gly His Gln Ser Thr Val Glu Asp Ile Ile Trp Pro Ile Val
 1               5                  10                  15
Arg Asn Asp Leu Ser Tyr Val Ser Gln His Trp Asn Glu Thr Thr Phe
            20                  25                  30
```

<210> 591
<211> 213
<212> PRT

<213> Thermomyces lanuginosus ATCC 200065

<400> 591

```
Asp Ile Trp Glu Glu Val His Ser Ser Ser Phe Phe Thr Thr Ala Val
 1               5                  10                  15
Gln Tyr Arg Ala Leu Val Gln Gly Ser Ala Leu Ala Ser Lys Leu Gly
            20                  25                  30
His Thr Cys Asp Asn Cys Gly Ser Gln Ala Pro Gln Ile Leu Cys Phe
        35                  40                  45
Leu Gln Ser Tyr Trp Thr Gly Ser His Ile Leu Ala Asn Thr Gly Gly
    50                  55                  60
Gly Arg Ser Gly Lys Asp Val Ser Thr Ile Leu Gly Val Ile Gly Ser
65                  70                  75                  80
Phe Asp Pro Asn Ala Asp Cys Asp Asp Val Thr Phe Gln Pro Cys Ser
                85                  90                  95
Ala Arg Ala Leu Ala Asn His Lys Gln Val Val Asp Ser Phe Arg Ser
            100                 105                 110
Ile Tyr Ala Ile Asn Ala Gly Ile Pro Ser Gly Ser Ala Val Ala Val
        115                 120                 125
Gly Arg Tyr Pro Glu Asp Val Tyr Gln Gly Gly His Pro Trp Tyr Leu
    130                 135                 140
Thr Thr Ala Ala Ala Ala Glu Gln Leu Tyr Asp Ala Ile Tyr Gln Trp
145                 150                 155                 160
Asn His Val Gly His Ile Asp Ile Asn Ala Val Asn Leu Asp Phe Phe
                165                 170                 175
Lys Ser Ile Tyr Pro Ser Ala Ala Glu Gly Thr Tyr Thr Ser Asp Ser
            180                 185                 190
Ser Thr Phe Gln Asp Ile Ile Ser Ala Val Arg Thr Tyr Ala Asp Gly
        195                 200                 205
Phe Leu Ser Val Ile
    210
```

<210> 592
<211> 197
<212> PRT
<213> Thermomyces lanuginosus ATCC 200065

<400> 592

```
Glu Lys Tyr Thr Pro Pro Asp Asn Leu Leu Ala Glu Gln Phe His Arg
 1               5                  10                  15
Glu Thr Gly Ile Pro Leu Ser Ala Ala Ser Leu Thr Trp Ser Tyr Ala
            20                  25                  30
Ala Leu Asn Thr Ala Ala Gln Arg Arg Ala Ser Ile Val Pro Ser Pro
        35                  40                  45
Trp Asn Ser Asn Ser Thr Asp Leu Pro Asp Lys Cys Ser Ala Thr Ser
    50                  55                  60
Ala Thr Gly Pro Tyr Ala Thr Pro Thr Asn Thr Ala Trp Pro Thr Thr
65                  70                  75                  80
Thr Gln Pro Pro Glu Arg Pro Ala Cys Thr Pro Pro Ser Glu Val Thr
                85                  90                  95
Leu Thr Phe Asn Ala Leu Val Asp Thr Ala Phe Gly Gln Asn Ile Tyr
            100                 105                 110
Leu Val Gly Ser Ile Pro Glu Leu Gly Ser Trp Asp Pro Ala Asn Ala
        115                 120                 125
Leu Leu Met Ser Ala Lys Ser Trp Thr Ser Gly Asn Pro Val Trp Thr
    130                 135                 140
Leu Ser Ile Ser Leu Pro Ala Gly Thr Ser Phe Glu Tyr Lys Phe Ile
145                 150                 155                 160
Arg Lys Asp Asp Gly Ser Ser Asp Val Val Trp Glu Ser Asp Pro Asn
                165                 170                 175
Arg Ser Tyr Asn Val Pro Lys Asp Cys Gly Ala Asn Thr Ala Thr Val
            180                 185                 190
Asn Ser Trp Trp Arg
        195
```

<210> 593
<211> 1854
<212> DNA

<213> Thermomyces lanuginosus ATCC 200065

<400> 593
atgttattcc aaccgacttt gtgcgcggcc cttggactcg ccgccttgat cgtccaaggc 60
ggagaagcca gacctgaaac aaccgtccca catgcaacgg gctcgctcga cgacttcctc 120
gccgcacaga gtccgattgc tttccaaggc atcctgaaca atatcgggcc tagcggagcg 180
tactcggaag gtgtcaatcc gggtgtggtc attgcgagtc caagtaaaca agatcccgac 240
tacttttaca cctgggtgcg cgacgctgct ctcactgtcc aatatctggt ggaggagctg 300
gttgcaggaa atgccagtct tcagttcctc attcaggact acatcagctc ccaggcacga 360
ctgcagacgg tggaaaatcc atccggctcc ctctcgtcgg gtggtctagg agagcccaag 420
tttcatgtcg acgagaccgc ctttacggac tcctggggcc gaccacagcg ggacggcccg 480
cctctccgcg ccattgccat gatttcgttt gccaattacc tgattgacaa cggtcatcaa 540
tcgactgtgg aggacatcat ctggccgatt gttcgcaatg acttgtccta tgtctcgcag 600
cattggaacg aaacaacttt tgacatctgg gaggaagtcc atagctcatc gtttttcacc 660
acggctgtcc agtaccgtgc tctggtccaa ggcagtgcct tggctagcaa gctcggccat 720
acctgcgaca actgcgggtc ccaagcaccg cagatccttt gcttcctgca gtcgtattgg 780
accgggtcgc acatcttagc caacaccggt ggcggccgct cgggaaagga cgtcagcacg 840
atcctcggcg tcattggctc gtttgatccg aacgccgact gtgatgacgt taccttccag 900
ccctgctcgg cccgggctct tgcaaatcac aagcaggtcg ttgacagctt ccgcagtatc 960
tatgccatca acgctggcat cccgtcaggg tcggctgttg cggttggacg ttatcccgag 1020
gatgtctatc agggtggaca cccctggtac ctaacaacgg ctgcggcggc ggagcagctt 1080
tacgacgcca tttaccagtg gaaccatgta gggcacatcg acatcaatgc tgtcaatctg 1140
gacttcttca agagcattta tccgtcagcc gccgagggca catacacatc agactcttca 1200
acatttcaag acattatatc tgctgtacgg acctatgcgg acgggtttct cagcgtaatt 1260
gagaaataca ctccgccgga taacttgctt gccgagcagt tccaccggga gacgggcatt 1320
ccactatcgg cagcttctct gacatggtcc tacgccgcgc tcaacacggc cgcgcagcgg 1380
cgagcgtcaa tcgtgccctc accgtggaac tctaacagca cagatctccc ggacaaatgc 1440
tcggcaacct cggcaacagg gccgtatgcc acgcccacaa acacggcatg gccaaccact 1500
acgcagccac cggagcggcc ggcatgcaca ccgccgtcgg aagtaacact caccttcaac 1560
gcgctcgtcg acaccgcgtt tggccagaat atttatctcg tgggctccat tccggagctc 1620
ggatcgtggg atccggccaa cgccctcttg atgagcgcaa agagctggac tagcggaaat 1680
ccggtctgga cgctatccat ttcccttcca gcaggaacct cttttgagta caagttcatt 1740
cgaaaggatg atggttcctc ggatgttgtc tgggaaagtg acccgaatcg ttcgtacaac 1800
gtgccgaagg attgcggtgc caacacggcc accgtgaatt cttggtggcg atga 1854

<210> 594
<211> 617
<212> PRT
<213> Thermomyces lanuginosus ATCC 200065

<220>
<221> SIGNAL
<222> (1)...(23)

<400> 594
Met Leu Phe Gln Pro Thr Leu Cys Ala Ala Leu Gly Leu Ala Ala Leu
 1   5   10   15
Ile Val Gln Gly Gly Glu Ala Arg Pro Glu Thr Thr Val Pro His Ala
            20   25   30
Thr Gly Ser Leu Asp Asp Phe Leu Ala Ala Gln Ser Pro Ile Ala Phe
        35   40   45
Gln Gly Ile Leu Asn Asn Ile Gly Pro Ser Gly Ala Tyr Ser Glu Gly
    50   55   60
Val Asn Pro Gly Val Val Ile Ala Ser Pro Ser Lys Gln Asp Pro Asp
65   70   75   80
Tyr Phe Tyr Thr Trp Val Arg Asp Ala Ala Leu Thr Val Gln Tyr Leu
                85   90   95
Val Glu Glu Leu Val Ala Gly Asn Ala Ser Leu Gln Phe Leu Ile Gln
            100   105   110
Asp Tyr Ile Ser Ser Gln Ala Arg Leu Gln Thr Val Glu Asn Pro Ser
        115   120   125
Gly Ser Leu Ser Ser Gly Gly Leu Gly Glu Pro Lys Phe His Val Asp
    130   135   140
Glu Thr Ala Phe Thr Asp Ser Trp Gly Arg Pro Gln Arg Asp Gly Pro
145   150   155   160
Pro Leu Arg Ala Ile Ala Met Ile Ser Phe Ala Asn Tyr Leu Ile Asp
                165   170   175
Asn Gly His Gln Ser Thr Val Glu Asp Ile Ile Trp Pro Ile Val Arg

```
            180                 185                 190
Asn Asp Leu Ser Tyr Val Ser Gln His Trp Asn Glu Thr Thr Phe Asp
        195                 200                 205
Ile Trp Glu Glu Val His Ser Ser Ser Phe Phe Thr Thr Ala Val Gln
    210                 215                 220
Tyr Arg Ala Leu Val Gln Gly Ser Ala Leu Ala Ser Lys Leu Gly His
225                 230                 235                 240
Thr Cys Asp Asn Cys Gly Ser Gln Ala Pro Gln Ile Leu Cys Phe Leu
                245                 250                 255
Gln Ser Tyr Trp Thr Gly Ser His Ile Leu Ala Asn Thr Gly Gly Gly
            260                 265                 270
Arg Ser Gly Lys Asp Val Ser Thr Ile Leu Gly Val Ile Gly Ser Phe
        275                 280                 285
Asp Pro Asn Ala Asp Cys Asp Asp Val Thr Phe Gln Pro Cys Ser Ala
    290                 295                 300
Arg Ala Leu Ala Asn His Lys Gln Val Val Asp Ser Phe Arg Ser Ile
305                 310                 315                 320
Tyr Ala Ile Asn Ala Gly Ile Pro Ser Gly Ser Ala Val Ala Val Gly
                325                 330                 335
Arg Tyr Pro Glu Asp Val Tyr Gln Gly Gly His Pro Trp Tyr Leu Thr
            340                 345                 350
Thr Ala Ala Ala Ala Glu Gln Leu Tyr Asp Ala Ile Tyr Gln Trp Asn
        355                 360                 365
His Val Gly His Ile Asp Ile Asn Ala Val Asn Leu Asp Phe Phe Lys
    370                 375                 380
Ser Ile Tyr Pro Ser Ala Ala Glu Gly Thr Tyr Thr Ser Asp Ser Ser
385                 390                 395                 400
Thr Phe Gln Asp Ile Ile Ser Ala Val Arg Thr Tyr Ala Asp Gly Phe
                405                 410                 415
Leu Ser Val Ile Glu Lys Tyr Thr Pro Pro Asp Asn Leu Leu Ala Glu
            420                 425                 430
Gln Phe His Arg Glu Thr Gly Ile Pro Leu Ser Ala Ala Ser Leu Thr
        435                 440                 445
Trp Ser Tyr Ala Ala Leu Asn Thr Ala Ala Gln Arg Arg Ala Ser Ile
    450                 455                 460
Val Pro Ser Pro Trp Asn Ser Asn Ser Thr Asp Leu Pro Asp Lys Cys
465                 470                 475                 480
Ser Ala Thr Ser Ala Thr Gly Pro Tyr Ala Thr Pro Thr Asn Thr Ala
                485                 490                 495
Trp Pro Thr Thr Thr Gln Pro Pro Glu Arg Pro Ala Cys Thr Pro Pro
            500                 505                 510
Ser Glu Val Thr Leu Thr Phe Asn Ala Leu Val Asp Thr Ala Phe Gly
        515                 520                 525
Gln Asn Ile Tyr Leu Val Gly Ser Ile Pro Glu Leu Gly Ser Trp Asp
    530                 535                 540
Pro Ala Asn Ala Leu Leu Met Ser Ala Lys Ser Trp Thr Ser Gly Asn
545                 550                 555                 560
Pro Val Trp Thr Leu Ser Ile Ser Leu Pro Ala Gly Thr Ser Phe Glu
                565                 570                 575
Tyr Lys Phe Ile Arg Lys Asp Asp Gly Ser Ser Asp Val Val Trp Glu
            580                 585                 590
Ser Asp Pro Asn Arg Ser Tyr Asn Val Pro Lys Asp Cys Gly Ala Asn
        595                 600                 605
Thr Ala Thr Val Asn Ser Trp Trp Arg
    610                 615

<210> 595
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 595
gcgaccttgg attcatggtt gagcaac                                          27

<210> 596
```

<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 596
cacaatagag acgaagccat cggcgaa 27

<210> 597
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 597
atgttattcc aaccgacttt gtgcgc 26

<210> 598
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 598
tcatcgccac caagaattca cggtg 25

<210> 599
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 599
gtctcgagaa aagagcaacg ggctcgctcg ac 32

<210> 600
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 600
gttctagatc atcgccacca agaattcacg gt 32

<210> 601
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 601
gtccatggca acgggctcgc tcgac 25

<210> 602
<211> 32
<212> DNA

<213> Artificial Sequence

<220>
<223> Primer

<400> 602
gttctagatc atcgccacca agaattcacg gt 32

<210> 603
<211> 1770
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 603
atgcaaaaca cagcgaaaaa ctccatctgg cagagggtgc gccacagcgc cattgcctta 60
tccgctctca gcttatcctt tggcctgcag gccagcgagt taccacaaat tccaccacag 120
caggtgaaca acaccatgta ccaggcattt tattgggatg cctaccctgg cctttgggcc 180
aatttaccgg ccatggcggc ccctttggcc gagcgtggca ttacctcgat gtggttgccg 240
ccggcggcaa aaggcatgaa tggcaccaac agcgttggct acgatgtcta tgacttctgg 300
gatctgggcg aattctacca aaaagggacc atcgccaccc gttacggcac ccgccagcag 360
ctggaccagg cgctgagtgc gctggatcaa ctgggtgttc aggcctattt tgatgtggtc 420
tttaaccacc gcatgggcgc cgatgcccag gaatacattc ctggctatgg actggcctgg 480
accgagtatc atctgcaagg tcggcaggcg cattataccc agcaaaactg gggctacttg 540
tggcacgact ttgactggaa ctggaccgcg tttaatggct ccgacaatca gctctacccc 600
ggcaaatggt ggggcaatac cttccacttc ccttatttga tgggtgagga tgtcgattac 660
aaccgctttg aagtgcagca ggaaatgaaa gcctgggggg agtggatcat caacagcgtt 720
ggctttagcg gctttcggat ggatgccatt gcccatgttg ataccgattt tacccgtgac 780
tggatcaatc acgtacagtg ggccaccagc gaggatgtgt tctttgtcgc agaagcctgg 840
gtcagtgata tcaacggcta tctggatgca gtcaatacgc cgcatttgcg cgcttttgat 900
ttcaatttgc gcgaagactt tgttgcttta agcagtggtg gcaaagacat gcgctggtgg 960
ggtggtctgg ttaatagcca gcaccgtgat cgggcggtca cttttgtcga taaccacgat 1020
accagccggg ccggcaaccc ttatggcatg ccgcaggtga tcaactacaa gaaccaggcc 1080
tacgcttaca ttctgttgcg tgagcatggg gtgccgactg tgtttgcccg cgattacgac 1140
gaatttggca tgacgccaac gctggataaa ttgattgagg cgcgccgcta ctttgcctac 1200
ggtcctggtc atgagtactc cggcaatacc gaggccgttt acgcctatgt gcgcgagggg 1260
cttagcactg tgccgggcac cggtctggtg atgctgatat cgggtcgaaa ctggggtggt 1320
cagcagtcgt tcaccatcaa cagccatcag ccgaatacca ccttttacga ttataccggc 1380
aatgtcagcg gcacggtgac caccaatgcg cagggttatg gcagcttccc ggtcactatg 1440
acggaaagta ccggttggtc agtctgggta ccacaatcca gtggtggcac tcagccggga 1500
tccattaccc tgcggatgac caaggatgtt ggctatggtt tttcactgtt cttcactggc 1560
agcagtgcag agctcaccaa ttggggcggc ggtattgaag gcacctggac cagtggccat 1620
gtctgggaag tgaccatccc ggatccgggc aactttgaat ggaaaacccg taaaggccca 1680
accggcggca gtggtcagga ctgggaaagt ggtagcaacc acaatcagac caatttgcac 1740
cccagtttta atggtgggtt tagatctcgc 1770

<210> 604
<211> 590
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(31)

<221> DOMAIN
<222> (44)...(395)
<223> Alpha amylase, catalytic domain

<400> 604
Met Gln Asn Thr Ala Lys Asn Ser Ile Trp Gln Arg Val Arg His Ser
 1 5 10 15
Ala Ile Ala Leu Ser Ala Leu Ser Leu Ser Phe Gly Leu Gln Ala Ser
 20 25 30

```
Glu Leu Pro Gln Ile Pro Pro Gln Gln Val Asn Asn Thr Met Tyr Gln
        35                  40                  45
Ala Phe Tyr Trp Asp Ala Tyr Pro Gly Leu Trp Ala Asn Leu Pro Ala
    50                  55                  60
Met Ala Ala Pro Leu Ala Glu Arg Gly Ile Thr Ser Met Trp Leu Pro
65                  70                  75                  80
Pro Ala Ala Lys Gly Met Asn Gly Thr Asn Ser Val Gly Tyr Asp Val
                85                  90                  95
Tyr Asp Phe Trp Asp Leu Gly Glu Phe Tyr Gln Lys Gly Thr Ile Ala
            100                 105                 110
Thr Arg Tyr Gly Thr Arg Gln Gln Leu Asp Gln Ala Leu Ser Ala Leu
        115                 120                 125
Asp Gln Leu Gly Val Gln Ala Tyr Phe Asp Val Val Phe Asn His Arg
    130                 135                 140
Met Gly Ala Asp Ala Gln Glu Tyr Ile Pro Gly Tyr Gly Leu Ala Trp
145                 150                 155                 160
Thr Glu Tyr His Leu Gln Gly Arg Gln Ala His Tyr Thr Gln Gln Asn
                165                 170                 175
Trp Gly Tyr Leu Trp His Asp Phe Asp Trp Asn Trp Thr Ala Phe Asn
            180                 185                 190
Gly Ser Asp Asn Gln Leu Tyr Pro Gly Lys Trp Trp Gly Asn Thr Phe
        195                 200                 205
His Phe Pro Tyr Leu Met Gly Glu Asp Val Asp Tyr Asn Arg Phe Glu
    210                 215                 220
Val Gln Gln Glu Met Lys Ala Trp Gly Glu Trp Ile Ile Asn Ser Val
225                 230                 235                 240
Gly Phe Ser Gly Phe Arg Met Asp Ala Ile Ala His Val Asp Thr Asp
                245                 250                 255
Phe Thr Arg Asp Trp Ile Asn His Val Gln Trp Ala Thr Ser Glu Asp
            260                 265                 270
Val Phe Phe Val Ala Glu Ala Trp Val Ser Asp Ile Asn Gly Tyr Leu
        275                 280                 285
Asp Ala Val Asn Thr Pro His Leu Arg Ala Phe Asp Phe Asn Leu Arg
    290                 295                 300
Glu Asp Phe Val Ala Leu Ser Ser Gly Gly Lys Asp Met Arg Trp Trp
305                 310                 315                 320
Gly Gly Leu Val Asn Ser Gln His Arg Asp Arg Ala Val Thr Phe Val
                325                 330                 335
Asp Asn His Asp Thr Ser Arg Ala Gly Asn Pro Tyr Gly Met Pro Gln
            340                 345                 350
Val Ile Asn Tyr Lys Asn Gln Ala Tyr Ala Tyr Ile Leu Leu Arg Glu
        355                 360                 365
His Gly Val Pro Thr Val Phe Ala Arg Asp Tyr Asp Glu Phe Gly Met
    370                 375                 380
Thr Pro Thr Leu Asp Lys Leu Ile Glu Ala Arg Arg Tyr Phe Ala Tyr
385                 390                 395                 400
Gly Pro Gly His Glu Tyr Ser Gly Asn Thr Glu Ala Val Tyr Ala Tyr
                405                 410                 415
Val Arg Glu Gly Leu Ser Thr Val Pro Gly Thr Gly Leu Val Met Leu
            420                 425                 430
Ile Ser Gly Arg Asn Trp Gly Gly Gln Gln Ser Phe Thr Ile Asn Ser
        435                 440                 445
His Gln Pro Asn Thr Thr Phe Tyr Asp Tyr Thr Gly Asn Val Ser Gly
    450                 455                 460
Thr Val Thr Thr Asn Ala Gln Gly Tyr Gly Ser Phe Pro Val Thr Met
465                 470                 475                 480
Thr Glu Ser Thr Gly Trp Ser Val Trp Val Pro Gln Ser Ser Gly Gly
                485                 490                 495
Thr Gln Pro Gly Ser Ile Thr Leu Arg Met Thr Lys Asp Val Gly Tyr
            500                 505                 510
Gly Phe Ser Leu Phe Phe Thr Gly Ser Ser Ala Glu Leu Thr Asn Trp
        515                 520                 525
Gly Gly Gly Ile Glu Gly Thr Trp Thr Ser Gly His Val Trp Glu Val
    530                 535                 540
Thr Ile Pro Asp Pro Gly Asn Phe Glu Trp Lys Thr Arg Lys Gly Pro
545                 550                 555                 560
Thr Gly Gly Ser Gly Gln Asp Trp Glu Ser Gly Ser Asn His Asn Gln
                565                 570                 575
```

Thr Asn Leu His Pro Ser Phe Asn Gly Gly Phe Arg Ser Arg
580 585 590

<210> 605
<211> 2160
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 605
atggttaacc atttgaaaaa atggatcgcc ggcatggcgt tgacgctcgc tctgttgacc 60
gggacggtcg ttcccggcct tccggtccag gtcgcttccg ccgcgcccga cacctccgta 120
tcgaacaagc agaacttcag cacggacgtg atttaccaga ttttcaccga ccggttcgcc 180
gacggcaatc ccgccaacaa tccgacgggc tcggcctaca gccccggctg cacgaatttg 240
cggctgtact gcggcggcga ctggaaaggc attacggaca agatcaacgc cggttatttc 300
accaatatgg gcgtaacggc gctgtggatt tcgcagccgg tggaaaacat ttattccgtc 360
atcaactact ccggcgtcaa caatacgtcc taccacgggt actgggctcg cgatttcaaa 420
aagaccaacc cttattacgg cacgttcgcg gatttgcaaa ccctgatcaa tgccgcgcat 480
gccaaaggca tcaagatcgt gatcgacttc gcgcccaatc acacctcgcc ggccctggaa 540
acggacagct ccttcgccga gaacggccgg ctgtacgaca acggcacgct gctcggcggc 600
tatacgaacg acacgaacgg ctacttccat cataacgggg gaaccgattt ctcctcgctg 660
gaagacggca tttacagaaa cctgtacgac ctcgccgacc tcaaccacaa cgtcagccgg 720
atcgacacct acttcaagga agcgatcaaa gtctggctcg acttgggaat cgacggcatc 780
cgcgtcgatg cggtcaagca tatgccgatg ggctggcaga agaactggat gtcctacatc 840
tacggttaca aaccggtgtt caccttcggc gagtggttcc tcggcacgaa tgagaacgac 900
ccgagcaacg tctatttcgc caacgaatcc ggcatgagcc tgctcgactt ccgtttcgcc 960
cagaaggttc gccaggtgct ccgcgacaac accgacacga tgtacggcct ggattcgatg 1020
atttcctcca ccgcctccga ctacgcgcag gtcgccgacc aggtcacctt catcgacaat 1080
cacgacatgg agcggttcaa atcgggctcg gtcagcaacc gccgggtgga gcaggcgctg 1140
gcgctgctgc tgacgtcccg cggcgttccc gccatctatt acggaacgga gcagtatatg 1200
agcggcggca ccgatccgga caaccgtgcg atgatgtcct ccttctccga gacgacgacc 1260
gcgtttaacg taatcaaaaa actcgctccg cttcgcaagt cgaatcccgc catcgcatac 1320
ggttcgaccc agcagcggtg gatcaacaac gacgtttaca tttacgagcg gaaattcgga 1380
aacagcgtgg cggtcgtcgc catcaaccgc aacttgtcca gctccgcgtc gatcagcgga 1440
ttggtcactt ccttcccggc gggcacctat acggaccaac tgggcggcct gctcggcggc 1500
ggcagcatca cggcaagcag caccggcgcc gtatcgacgt ttacgctcgg agcgggagcc 1560
gtagccgtat ggcaatatac gccggcccaa acgacgccgg tgatcggcca tgtcggtccg 1620
gtcatgggca aagccggcaa caccgtcacg atcgacggca gaggtttcgg ctccacgaag 1680
ggcaccgtct acttcggcac gaccgcggtg accggttcgg gcatcgtttc gtgggaagac 1740
acgcagatca aggtcgtcgt gccttccgtc gccgcaggca aatacaacgt gaaagtccgc 1800
acgtcgggct cggtcgacag caacagctac aacggcttta atgtgttgac gggcaaccag 1860
gtgtccgtgc gcttcatcgt gcaaaacgcg acgaccgtca gcggcgaaaa cgtctatctc 1920
gtcggcaacg tcgccgaact cggcaactgg gatccggcct cggccatagg cccgatgttc 1980
aaccgggtca tccgaaccta tccgacctgg tatttcgacg tcagcgtacc ggccggcacc 2040
gccctggaat tcaaatttat caaaaagaac ggctccgtcg tgacctggga aggcggctcc 2100
aaccatacgt acacgacccc ggccagcggc accggcacca taacggtcaa ctggcagtaa 2160

<210> 606
<211> 719
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(34)

<221> DOMAIN
<222> (53)...(432)
<223> Alpha amylase, catalytic domain

<221> DOMAIN
<222> (620)...(716)
<223> Starch binding domain

<221> DOMAIN
<222> (533)...(616)
<223> IPT/TIG domain

<221> DOMAIN
<222> (444)...(528)
<223> Alpha amylase, C-terminal all-beta domain

<400> 606

```
Met Val Asn His Leu Lys Lys Trp Ile Ala Gly Met Ala Leu Thr Leu
 1               5                  10                  15
Ala Leu Leu Thr Gly Thr Val Val Pro Gly Leu Pro Val Gln Val Ala
            20                  25                  30
Ser Ala Ala Pro Asp Thr Ser Val Ser Asn Lys Gln Asn Phe Ser Thr
        35                  40                  45
Asp Val Ile Tyr Gln Ile Phe Thr Asp Arg Phe Ala Asp Gly Asn Pro
    50                  55                  60
Ala Asn Asn Pro Thr Gly Ser Ala Tyr Ser Pro Gly Cys Thr Asn Leu
65                  70                  75                  80
Arg Leu Tyr Cys Gly Gly Asp Trp Lys Gly Ile Thr Asp Lys Ile Asn
                85                  90                  95
Ala Gly Tyr Phe Thr Asn Met Gly Val Thr Ala Leu Trp Ile Ser Gln
            100                 105                 110
Pro Val Glu Asn Ile Tyr Ser Val Ile Asn Tyr Ser Gly Val Asn Asn
        115                 120                 125
Thr Ser Tyr His Gly Tyr Trp Ala Arg Asp Phe Lys Lys Thr Asn Pro
    130                 135                 140
Tyr Tyr Gly Thr Phe Ala Asp Leu Gln Thr Leu Ile Asn Ala Ala His
145                 150                 155                 160
Ala Lys Gly Ile Lys Ile Val Ile Asp Phe Ala Pro Asn His Thr Ser
                165                 170                 175
Pro Ala Leu Glu Thr Asp Ser Ser Phe Ala Glu Asn Gly Arg Leu Tyr
            180                 185                 190
Asp Asn Gly Thr Leu Leu Gly Gly Tyr Thr Asn Asp Thr Asn Gly Tyr
        195                 200                 205
Phe His His Asn Gly Gly Thr Asp Phe Ser Ser Leu Glu Asp Gly Ile
    210                 215                 220
Tyr Arg Asn Leu Tyr Asp Leu Ala Asp Leu Asn His Asn Val Ser Arg
225                 230                 235                 240
Ile Asp Thr Tyr Phe Lys Glu Ala Ile Lys Val Trp Leu Asp Leu Gly
                245                 250                 255
Ile Asp Gly Ile Arg Val Asp Ala Val Lys His Met Pro Met Gly Trp
            260                 265                 270
Gln Lys Asn Trp Met Ser Tyr Ile Tyr Gly Tyr Lys Pro Val Phe Thr
        275                 280                 285
Phe Gly Glu Trp Phe Leu Gly Thr Asn Glu Asn Asp Pro Ser Asn Val
    290                 295                 300
Tyr Phe Ala Asn Glu Ser Gly Met Ser Leu Leu Asp Phe Arg Phe Ala
305                 310                 315                 320
Gln Lys Val Arg Gln Val Leu Arg Asp Asn Thr Asp Thr Met Tyr Gly
                325                 330                 335
Leu Asp Ser Met Ile Ser Ser Thr Ala Ser Asp Tyr Ala Gln Val Ala
            340                 345                 350
Asp Gln Val Thr Phe Ile Asp Asn His Asp Met Glu Arg Phe Lys Ser
        355                 360                 365
Gly Ser Val Ser Asn Arg Arg Val Glu Gln Ala Leu Ala Leu Leu Leu
    370                 375                 380
Thr Ser Arg Gly Val Pro Ala Ile Tyr Tyr Gly Thr Glu Gln Tyr Met
385                 390                 395                 400
Ser Gly Gly Thr Asp Pro Asp Asn Arg Ala Met Met Ser Ser Phe Ser
                405                 410                 415
Glu Thr Thr Thr Ala Phe Asn Val Ile Lys Lys Leu Ala Pro Leu Arg
            420                 425                 430
Lys Ser Asn Pro Ala Ile Ala Tyr Gly Ser Thr Gln Gln Arg Trp Ile
        435                 440                 445
Asn Asn Asp Val Tyr Ile Tyr Glu Arg Lys Phe Gly Asn Ser Val Ala
    450                 455                 460
Val Val Ala Ile Asn Arg Asn Leu Ser Ser Ser Ala Ser Ile Ser Gly
```

```
465                 470                 475                 480
Leu Val Thr Ser Phe Pro Ala Gly Thr Tyr Thr Asp Gln Leu Gly Gly
                485                 490                 495
Leu Leu Gly Gly Gly Ser Ile Thr Ala Ser Ser Thr Gly Ala Val Ser
            500                 505                 510
Thr Phe Thr Leu Gly Ala Gly Ala Val Ala Val Trp Gln Tyr Thr Pro
        515                 520                 525
Ala Gln Thr Thr Pro Val Ile Gly His Val Gly Pro Val Met Gly Lys
    530                 535                 540
Ala Gly Asn Thr Val Thr Ile Asp Gly Arg Gly Phe Gly Ser Thr Lys
545                 550                 555                 560
Gly Thr Val Tyr Phe Gly Thr Thr Ala Val Thr Gly Ser Gly Ile Val
                565                 570                 575
Ser Trp Glu Asp Thr Gln Ile Lys Val Val Val Pro Ser Val Ala Ala
            580                 585                 590
Gly Lys Tyr Asn Val Lys Val Arg Thr Ser Gly Ser Val Asp Ser Asn
        595                 600                 605
Ser Tyr Asn Gly Phe Asn Val Leu Thr Gly Asn Gln Val Ser Val Arg
    610                 615                 620
Phe Ile Val Gln Asn Ala Thr Thr Val Ser Gly Glu Asn Val Tyr Leu
625                 630                 635                 640
Val Gly Asn Val Ala Glu Leu Gly Asn Trp Asp Pro Ala Ser Ala Ile
                645                 650                 655
Gly Pro Met Phe Asn Arg Val Ile Arg Thr Tyr Pro Thr Trp Tyr Phe
            660                 665                 670
Asp Val Ser Val Pro Ala Gly Thr Ala Leu Glu Phe Lys Phe Ile Lys
        675                 680                 685
Lys Asn Gly Ser Val Val Thr Trp Glu Gly Gly Ser Asn His Thr Tyr
    690                 695                 700
Thr Thr Pro Ala Ser Gly Thr Gly Thr Ile Thr Val Asn Trp Gln
705                 710                 715
```

<210> 607
<211> 3705
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 607

```
ttggaaacga cgatcattac attcttcttc aatgcgaaaa caaaagaggt gtacaccgac      60
tacaacccca atggctctga cggtgttgtt cagaaagacc gattaaaaca caatacatgg     120
gatacgctgt accgccagcc gtttggtgcg gtgaaagcgg gaacgaaggt gacgcttcgt     180
ctggcggcga aaaaaggcga tttgacgaag gcggatgtgt atgtgaaaaa tgcgacgacc     240
ggaaccgcca aagtctacac gatggaaaaa gtcggcgtgc tcggggatga tgaatattgg     300
gaggctgcgt ttacccttc tgcgccgggg gtatacggtt ataaatttat cgctgtcgat     360
gccgggatga aggcggaata tggtgaagat acgaaagaag gacaatgggg gagggcggtc     420
gacaaaaatg cggggctgtt tcagctgacc gtctatgatc cgaattacaa aacaccagat     480
tggatgaaag aagcggttgt gtatcaaatt ttcccggatc ggttctttaa tggcaaccct     540
tcaaatgata acagcaagca gcaggcacgc ggggcgcagc cgattgagca tcgcgattgg     600
tcggatttgc ccgataatcc gcgcctgaaa gggacgagcg gctacgatgg cgacggtgaa     660
tggtcgaatg acttttcgg cggagacatc gccggaattg aacaaaagtt ggattatttg     720
cagtcgcttg gagtgaacac gatttactta aatccgatcg ccaatgcgcc atcgaaccat     780
aaatatgatg cgagcaatta caaagaattg gatccgatgt tcggttcccc ggaagaattc     840
caatcgtttg tgcaggcgct tgcgaaccgg gggatgcatc tcatcttaga cggggtgttc     900
aaccacgtat ccgacgattc gatttacttt gaccgctacc accgctatcc gaccgtcggt     960
gcgtatgaat attgggaagc ggtttacgat ttgatgaatg aaaaaggatt gagcgaggaa    1020
gaagcgcgga aacaagtgga agagaagttc aaacaagagg gacagacgtt cagcccgtat    1080
gggtttcacc tttggttcaa tattgaaaac aaaaaagtcg atggccatta tcaataccaa    1140
tcatggtggg gctatgacag tctgccggag tttaagtcgg tgacggggga aaaagtgccg    1200
catccgagtg aattgaacaa cgatgcgctc gcgaattaca ttttccgtga atcggattcg    1260
gtggcgaaaa gctggattgc cctcggcgcc tccggctggc ggttggatgt ggccaatgag    1320
gtggatccgg cgttttggcg cgagtttcgc caagaattgc ttcaagggtc gtacggtcgt    1380
ggtccgacgt taaaagaggg ggagcagccg ctcattttag gggaaatttg ggatgacgca    1440
tcgaaatatt ttctaggcga ccagtacgat tccgtgatga actaccggtt ccgcggggcg    1500
gtgcttgact ttttgaaaaa cggaaatgca gaagaggcgg acaagcggct gacggccata    1560
agggaagact acccaagtga agcgttttat gcgctgatga acttaatcgg ttcgcatgac    1620
```

```
acggcgcggg cggtctttct gcttgggaac ggaacggatt catccgagcg ggcggagctt    1680
gatccgaatt ataatgagga acttgggaaa aagcggctca agctggcggt gattttgcag    1740
atgggatacc cgggagcgcc gacgatttat tacggcgatg aagcgggagt aacaggctca    1800
agagacccag acaaccgccg cacgtatccg tggggcaaag aagatcaaaa tctgttgtcc    1860
cattatcaga aagtggggca cattcgccag caccatcaat cgttgttggc ccatggcgac    1920
atcaagacgg tgtatgcgca aggggatgta tacgtatttg cccgccaata cgggcgtgaa    1980
gtggcgctca ttgccctcaa ccgcggcaat gaggacaaga cggtggcgct tgacgtcgct    2040
tcgttgcttc cgaacggcac cgtgcttacg gatgagttgc atgatggcgg ggaagctacg    2100
gtcgctggcg gaacgttgac ggtcacgatt ccggccctgg atggacggat gatgtttggg    2160
acggtgacgg cggaaatgcc ggcagcagtc agcaatttgc aggcgagcgc ttcggatggc    2220
tgcgtgacgt taacgtggga aggaaatgcg tcgagatacc gaatttacga gtccacgtta    2280
aaaggtgccg gttatacgat ggtgcaagag acggaaacaa cttcggccac gatcggttcg    2340
ttgacgaacg gaacagccta ttactttgcc gttgcggcgg tcgatgaaaa cgggaatgaa    2400
tcaccgaagg tcgaaacgaa tcgcgtcgtt cctcattacc cgctgacgag cgacaatgtc    2460
cagttcgtga caacgttaag cgatgccaca ctggatttgt caaagccgca gcaagtggat    2520
gtccatgtca acatcgacaa tgtgacaagc aaaggagcag ctgatgggtt gcaagcggtg    2580
ttgcaagtga aaggcccgca tgacgaaaca tggaaagaat acagagcggc ttaccaagga    2640
caagacggcg acgccaacgt gttccgagct gccttcactc cgctcgccgc agggacgtat    2700
acgtatcgtt atgcgctgac gaccaacctc ggcgaggagt gggtgtatac agaagagaag    2760
caagtgacgt ttgcggcaga caacagcgac caaatagcgc cagcagacgc catcgagctg    2820
cggcagcctg cggttgaatc cggacaagtg aatttatcat ggacgtttgt tgggaaaaaa    2880
gatggggatg cttatttgtt agccatcgag cgcaacggtg atatcgtgca tacaaccact    2940
tcgatcggcg attcatttac agactacgat gtcgaaaacg gcaccgagta cacgtatgtt    3000
gtcaagttgt atgaccgcgc cggcaatgtt gtggcgtcaa acacggtcaa ggtgacgccg    3060
gacattgtga tggtgaaagt gatttttaaa gtgagggcgc cggattacac accgttggat    3120
gcccgaatta cgattccgaa cagcttgaac ggctggaaca caggggcctg ggagatgtcg    3180
cgcaacggtg cggtgacgcc cgattggcaa tttaccgtcg aggtgcagga aggggaaacg    3240
atcacctata agtatgtgaa aggcggatcg tgggatcaag aggggttggc cgaccatacg    3300
cgtgaggacg acaacgatga tgacgtgagc tactacggct atgggacgat tggcaccgac    3360
ttgaaagtga cggtccacaa tgaaggaaac aatacgatga ttgtgcaaga ccgcattttg    3420
cgctggatcg atatgccggt cgtcatcgaa gaggtgcaaa aacaaggaag tcaagtgacg    3480
atcaagggca atgccattaa aaacggtgtt ttgacgatca atggcgagcg ggtgccgatt    3540
gatggccgga tggcattctc gtacacgttt gcgccggcca gccatcaaaa agaagtgttg    3600
atccatatcg aaccatcggc cgaaagcaaa acagccattt tcaacaacga cggcggagcg    3660
attgcgaaaa acacaaaaga ttacgtgctg aatttagaaa cgaag                    3705

<210> 608
<211> 1235
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> DOMAIN
<222> (169)...(629)
<223> Alpha amylase, catalytic domain

<221> DOMAIN
<222> (727)...(801)
<223> Fibronectin type III domain

<221> DOMAIN
<222> (646)...(724)
<223> Alpha amylase, C-terminal all-beta domain

<400> 608
Met Glu Thr Thr Ile Ile Thr Phe Phe Phe Asn Ala Lys Thr Lys Glu
 1               5                  10                  15
Val Tyr Thr Asp Tyr Asn Pro Asn Gly Ser Asp Gly Val Val Gln Lys
            20                  25                  30
Asp Arg Leu Lys His Asn Thr Trp Asp Thr Leu Tyr Arg Gln Pro Phe
        35                  40                  45
Gly Ala Val Lys Ala Gly Thr Lys Val Thr Leu Arg Leu Ala Ala Lys
    50                  55                  60
Lys Gly Asp Leu Thr Lys Ala Asp Val Tyr Val Lys Asn Ala Thr Thr
65                  70                  75                  80
Gly Thr Ala Lys Val Tyr Thr Met Glu Lys Val Gly Val Leu Gly Asp
```

```
                85                  90                  95
Asp Glu Tyr Trp Glu Ala Ala Phe Thr Pro Ser Ala Pro Gly Val Tyr
            100                 105                 110
Gly Tyr Lys Phe Ile Ala Val Asp Ala Gly Met Lys Ala Glu Tyr Gly
        115                 120                 125
Glu Asp Thr Lys Glu Gly Gln Trp Gly Arg Ala Val Asp Lys Asn Ala
    130                 135                 140
Gly Leu Phe Gln Leu Thr Val Tyr Asp Pro Asn Tyr Lys Thr Pro Asp
145                 150                 155                 160
Trp Met Lys Glu Ala Val Val Tyr Gln Ile Phe Pro Asp Arg Phe Phe
                165                 170                 175
Asn Gly Asn Pro Ser Asn Asp Asn Ser Lys Gln Gln Ala Arg Gly Ala
            180                 185                 190
Gln Pro Ile Glu His Arg Asp Trp Ser Asp Leu Pro Asp Asn Pro Arg
        195                 200                 205
Leu Lys Gly Thr Ser Gly Tyr Asp Gly Asp Gly Glu Trp Ser Asn Asp
    210                 215                 220
Phe Phe Gly Gly Asp Ile Ala Gly Ile Glu Gln Lys Leu Asp Tyr Leu
225                 230                 235                 240
Gln Ser Leu Gly Val Asn Thr Ile Tyr Leu Asn Pro Ile Ala Asn Ala
                245                 250                 255
Pro Ser Asn His Lys Tyr Asp Ala Ser Asn Tyr Lys Glu Leu Asp Pro
            260                 265                 270
Met Phe Gly Ser Pro Glu Glu Phe Gln Ser Phe Val Gln Ala Leu Ala
        275                 280                 285
Asn Arg Gly Met His Leu Ile Leu Asp Gly Val Phe Asn His Val Ser
    290                 295                 300
Asp Asp Ser Ile Tyr Phe Asp Arg Tyr His Arg Tyr Pro Thr Val Gly
305                 310                 315                 320
Ala Tyr Glu Tyr Trp Glu Ala Val Tyr Asp Leu Met Asn Glu Lys Gly
                325                 330                 335
Leu Ser Glu Glu Glu Ala Arg Lys Gln Val Glu Glu Lys Phe Lys Gln
            340                 345                 350
Glu Gly Gln Thr Phe Ser Pro Tyr Gly Phe His Leu Trp Phe Asn Ile
        355                 360                 365
Glu Asn Lys Lys Val Asp Gly His Tyr Gln Tyr Gln Ser Trp Trp Gly
    370                 375                 380
Tyr Asp Ser Leu Pro Glu Phe Lys Ser Val Thr Gly Glu Lys Val Pro
385                 390                 395                 400
His Pro Ser Glu Leu Asn Asn Asp Ala Leu Ala Asn Tyr Ile Phe Arg
                405                 410                 415
Glu Ser Asp Ser Val Ala Lys Ser Trp Ile Ala Leu Gly Ala Ser Gly
            420                 425                 430
Trp Arg Leu Asp Val Ala Asn Glu Val Asp Pro Ala Phe Trp Arg Glu
        435                 440                 445
Phe Arg Gln Glu Leu Leu Gln Gly Ser Tyr Gly Arg Gly Pro Thr Leu
    450                 455                 460
Lys Glu Gly Glu Gln Pro Leu Ile Leu Gly Glu Ile Trp Asp Asp Ala
465                 470                 475                 480
Ser Lys Tyr Phe Leu Gly Asp Gln Tyr Asp Ser Val Met Asn Tyr Arg
                485                 490                 495
Phe Arg Gly Ala Val Leu Asp Phe Leu Lys Asn Gly Asn Ala Glu Glu
            500                 505                 510
Ala Asp Lys Arg Leu Thr Ala Ile Arg Glu Asp Tyr Pro Ser Glu Ala
        515                 520                 525
Phe Tyr Ala Leu Met Asn Leu Ile Gly Ser His Asp Thr Ala Arg Ala
    530                 535                 540
Val Phe Leu Leu Gly Asn Gly Thr Asp Ser Ser Glu Arg Ala Glu Leu
545                 550                 555                 560
Asp Pro Asn Tyr Asn Glu Glu Leu Gly Lys Lys Arg Leu Lys Leu Ala
                565                 570                 575
Val Ile Leu Gln Met Gly Tyr Pro Gly Ala Pro Thr Ile Tyr Tyr Gly
            580                 585                 590
Asp Glu Ala Gly Val Thr Gly Ser Arg Asp Pro Asp Asn Arg Arg Thr
        595                 600                 605
Tyr Pro Trp Gly Lys Glu Asp Gln Asn Leu Leu Ser His Tyr Gln Lys
    610                 615                 620
Val Gly His Ile Arg Gln His His Gln Ser Leu Leu Ala His Gly Asp
```

625 630 635 640
Ile Lys Thr Val Tyr Ala Gln Gly Asp Val Tyr Val Phe Ala Arg Gln
645 650 655
Tyr Gly Arg Glu Val Ala Leu Ile Ala Leu Asn Arg Gly Asn Glu Asp
660 665 670
Lys Thr Val Ala Leu Asp Val Ala Ser Leu Leu Pro Asn Gly Thr Val
675 680 685
Leu Thr Asp Glu Leu His Asp Gly Gly Glu Ala Thr Val Ala Gly Gly
690 695 700
Thr Leu Thr Val Thr Ile Pro Ala Leu Asp Gly Arg Met Met Phe Gly
705 710 715 720
Thr Val Thr Ala Glu Met Pro Ala Ala Val Ser Asn Leu Gln Ala Ser
725 730 735
Ala Ser Asp Gly Cys Val Thr Leu Thr Trp Glu Gly Asn Ala Ser Arg
740 745 750
Tyr Arg Ile Tyr Glu Ser Thr Leu Lys Gly Ala Gly Tyr Thr Met Val
755 760 765
Gln Glu Thr Glu Thr Thr Ser Ala Thr Ile Gly Ser Leu Thr Asn Gly
770 775 780
Thr Ala Tyr Tyr Phe Ala Val Ala Ala Val Asp Glu Asn Gly Asn Glu
785 790 795 800
Ser Pro Lys Val Glu Thr Asn Arg Val Val Pro His Tyr Pro Leu Thr
805 810 815
Ser Asp Asn Val Gln Phe Val Thr Thr Leu Ser Asp Ala Thr Leu Asp
820 825 830
Leu Ser Lys Pro Gln Gln Val Asp Val His Val Asn Ile Asp Asn Val
835 840 845
Thr Ser Lys Gly Ala Ala Asp Gly Leu Gln Ala Val Leu Gln Val Lys
850 855 860
Gly Pro His Asp Glu Thr Trp Lys Glu Tyr Arg Ala Ala Tyr Gln Gly
865 870 875 880
Gln Asp Gly Asp Ala Asn Val Phe Arg Ala Ala Phe Thr Pro Leu Ala
885 890 895
Ala Gly Thr Tyr Thr Tyr Arg Tyr Ala Leu Thr Thr Asn Leu Gly Glu
900 905 910
Glu Trp Val Tyr Thr Glu Glu Lys Gln Val Thr Phe Ala Ala Asp Asn
915 920 925
Ser Asp Gln Ile Ala Pro Ala Asp Ala Ile Glu Leu Arg Gln Pro Ala
930 935 940
Val Glu Ser Gly Gln Val Asn Leu Ser Trp Thr Phe Val Gly Lys Lys
945 950 955 960
Asp Gly Asp Ala Tyr Leu Leu Ala Ile Glu Arg Asn Gly Asp Ile Val
965 970 975
His Thr Thr Thr Ser Ile Gly Asp Ser Phe Thr Asp Tyr Asp Val Glu
980 985 990
Asn Gly Thr Glu Tyr Thr Tyr Val Val Lys Leu Tyr Asp Arg Ala Gly
995 1000 1005
Asn Val Val Ala Ser Asn Thr Val Lys Val Thr Pro Asp Ile Val Met
1010 1015 1020
Val Lys Val Ile Phe Lys Val Arg Ala Pro Asp Tyr Thr Pro Leu Asp
1025 1030 1035 1040
Ala Arg Ile Thr Ile Pro Asn Ser Leu Asn Gly Trp Asn Thr Gly Ala
1045 1050 1055
Trp Glu Met Ser Arg Asn Gly Ala Val Thr Pro Asp Trp Gln Phe Thr
1060 1065 1070
Val Glu Val Gln Glu Gly Glu Thr Ile Thr Tyr Lys Tyr Val Lys Gly
1075 1080 1085
Gly Ser Trp Asp Gln Glu Gly Leu Ala Asp His Thr Arg Glu Asp Asp
1090 1095 1100
Asn Asp Asp Asp Val Ser Tyr Tyr Gly Tyr Gly Thr Ile Gly Thr Asp
1105 1110 1115 1120
Leu Lys Val Thr Val His Asn Glu Gly Asn Asn Thr Met Ile Val Gln
1125 1130 1135
Asp Arg Ile Leu Arg Trp Ile Asp Met Pro Val Val Ile Glu Glu Val
1140 1145 1150
Gln Lys Gln Gly Ser Gln Val Thr Ile Lys Gly Asn Ala Ile Lys Asn
1155 1160 1165
Gly Val Leu Thr Ile Asn Gly Glu Arg Val Pro Ile Asp Gly Arg Met

```
    1170                1175                1180
Ala Phe Ser Tyr Thr Phe Ala Pro Ala Ser His Gln Lys Glu Val Leu
1185                1190                1195                1200
Ile His Ile Glu Pro Ser Ala Glu Ser Lys Thr Ala Ile Phe Asn Asn
                1205                1210                1215
Asp Gly Gly Ala Ile Ala Lys Asn Thr Lys Asp Tyr Val Leu Asn Leu
            1220                1225                1230
Glu Thr Lys
        1235
```

<210> 609
<211> 1764
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 609

```
atgcaaaaca cagcgaaaaa ctccatctgg cagagggtgc gccacagcgc cattgcctta      60
tccgctctca gcttatcctt tggcctgcag gccagcgagt taccacaaat tccaccacag     120
caggtgaaca acaccatgta ccaggcattt tattgggatg cctaccctgg cctttgggcc     180
aatttaccgg ccatggcggc ccctttggcc gagcgtggca ttacctcgat gtggttgccg     240
cccgccgcca aaggcatgaa tggtactttc agtgtcggtt acgatgtata cgatttctgg     300
gatctgggcg agtttaacca aaaaggcacc accgccaccc gttacggtac tcgtcagcag     360
ctgcaacaag cactgagtgc tctggaccaa ctgggtattc aggcctattt tgatgtggtg     420
tttaaccacc gcatgggcgc cgatgcacag gagaatattc ctggctttgg cctggcctgg     480
accgagtatc atctgcaagg tcgtcaggcg cattataccc agcaaaactg gggctacttg     540
tggcacgact ttgactggaa ctggaccgcg tttaatggct ccgacaatca gctctacccc     600
ggcaaatggt ggggcaatac cttccacttc ccttatttga tgggtgagga tgtcgattac     660
aaccgctttg aagtgcagca ggaaatgaaa gcctgggggg agtggatcat caacagcgtt     720
ggctttagcg gctttcggat ggatgccatt gcccatgttg ataccgattt tacccgtgac     780
tggatcaatc acgtacagtg ggccaccagc gaggatgtgt tctttgtcgc agaagcctgg     840
gtcagtgata taaacggcta tctggatgca gtcaatacgc cgcatttgcg cgcttttgat     900
ttcaatttgc gcgaagactt tgttgcttta agcagtggtg gcaaagacat gcgctggtgg     960
ggtggtctgg ttaatagcca gcaccgtgat cgggcggtca cttttgtcga taaccacgat    1020
accagccggg ccggcaaccc ttatggcatg ccgcaggtga tcaactacaa gaaccaggcc    1080
tacgcttaca ttctgttgcg tgagcatggg gtgccgactg tgtttgcccg cgattacgac    1140
gaatttggca tggcgccaac gctggataaa ttgattgaag cacggcgtta ctttgcctat    1200
ggtcctggcc atgagtactc cggcaatacc gaggccgtct acgcctgtgt gcgcgaaggg    1260
cttagcactg tgccgggcac cggtctggtg atgctgatat cgggtcgaaa ctggggtggt    1320
cagcagtcgt tcaccatcaa cagccatcag ccgaatacca ccttttacga ttataccggc    1380
aatgtcagtg gcactgtgac caccaatgcg cagggttatg gcagcttccc ggtcactatg    1440
acggaaagta ccggttggtc agtctgggta ccacaatcca gtggtggcac tcagccggga    1500
tccattaccc tgcggataac caaggatgtt ggctatggtt tttcactgtt cttcactggc    1560
agcagtgcag agctcactaa ttggggcggc ggtattgaag gcacctggac cagtggccat    1620
gtctgggaag tgaccatccc ggatcccggc aactttgaat ggaaaacccg taaaggccca    1680
accggtggca gtggtcagga ctgggaaagt gacagcaacc acaaccagac caatttgcac    1740
cccagtttta atggtgggtt ttaa                                           1764
```

<210> 610
<211> 587
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(31)

<221> DOMAIN
<222> (44)...(395)
<223> Alpha amylase, catalytic domain

<400> 610

```
Met Gln Asn Thr Ala Lys Asn Ser Ile Trp Gln Arg Val Arg His Ser
 1               5                  10                  15
```

```
Ala Ile Ala Leu Ser Ala Leu Ser Leu Ser Phe Gly Leu Gln Ala Ser
            20                  25                  30
Glu Leu Pro Gln Ile Pro Pro Gln Gln Val Asn Asn Thr Met Tyr Gln
        35                  40                  45
Ala Phe Tyr Trp Asp Ala Tyr Pro Gly Leu Trp Ala Asn Leu Pro Ala
    50                  55                  60
Met Ala Ala Pro Leu Ala Glu Arg Gly Ile Thr Ser Met Trp Leu Pro
65                  70                  75                  80
Pro Ala Ala Lys Gly Met Asn Gly Thr Phe Ser Val Gly Tyr Asp Val
                85                  90                  95
Tyr Asp Phe Trp Asp Leu Gly Glu Phe Asn Gln Lys Gly Thr Thr Ala
            100                 105                 110
Thr Arg Tyr Gly Thr Arg Gln Gln Leu Gln Gln Ala Leu Ser Ala Leu
        115                 120                 125
Asp Gln Leu Gly Ile Gln Ala Tyr Phe Asp Val Val Phe Asn His Arg
    130                 135                 140
Met Gly Ala Asp Ala Gln Glu Asn Ile Pro Gly Phe Gly Leu Ala Trp
145                 150                 155                 160
Thr Glu Tyr His Leu Gln Gly Arg Gln Ala His Tyr Thr Gln Gln Asn
                165                 170                 175
Trp Gly Tyr Leu Trp His Asp Phe Asp Trp Asn Trp Thr Ala Phe Asn
            180                 185                 190
Gly Ser Asp Asn Gln Leu Tyr Pro Gly Lys Trp Trp Gly Asn Thr Phe
        195                 200                 205
His Phe Pro Tyr Leu Met Gly Glu Asp Val Asp Tyr Asn Arg Phe Glu
    210                 215                 220
Val Gln Gln Glu Met Lys Ala Trp Gly Glu Trp Ile Ile Asn Ser Val
225                 230                 235                 240
Gly Phe Ser Gly Phe Arg Met Asp Ala Ile Ala His Val Asp Thr Asp
                245                 250                 255
Phe Thr Arg Asp Trp Ile Asn His Val Gln Trp Ala Thr Ser Glu Asp
            260                 265                 270
Val Phe Phe Val Ala Glu Ala Trp Val Ser Asp Ile Asn Gly Tyr Leu
        275                 280                 285
Asp Ala Val Asn Thr Pro His Leu Arg Ala Phe Asp Phe Asn Leu Arg
    290                 295                 300
Glu Asp Phe Val Ala Leu Ser Ser Gly Gly Lys Asp Met Arg Trp Trp
305                 310                 315                 320
Gly Gly Leu Val Asn Ser Gln His Arg Asp Arg Ala Val Thr Phe Val
                325                 330                 335
Asp Asn His Asp Thr Ser Arg Ala Gly Asn Pro Tyr Gly Met Pro Gln
            340                 345                 350
Val Ile Asn Tyr Lys Asn Gln Ala Tyr Ala Tyr Ile Leu Leu Arg Glu
        355                 360                 365
His Gly Val Pro Thr Val Phe Ala Arg Asp Tyr Asp Glu Phe Gly Met
    370                 375                 380
Ala Pro Thr Leu Asp Lys Leu Ile Glu Ala Arg Arg Tyr Phe Ala Tyr
385                 390                 395                 400
Gly Pro Gly His Glu Tyr Ser Gly Asn Thr Glu Ala Val Tyr Ala Cys
                405                 410                 415
Val Arg Glu Gly Leu Ser Thr Val Pro Gly Thr Gly Leu Val Met Leu
            420                 425                 430
Ile Ser Gly Arg Asn Trp Gly Gly Gln Gln Ser Phe Thr Ile Asn Ser
        435                 440                 445
His Gln Pro Asn Thr Thr Phe Tyr Asp Tyr Thr Gly Asn Val Ser Gly
    450                 455                 460
Thr Val Thr Thr Asn Ala Gln Gly Tyr Gly Ser Phe Pro Val Thr Met
465                 470                 475                 480
Thr Glu Ser Thr Gly Trp Ser Val Trp Val Pro Gln Ser Ser Gly Gly
                485                 490                 495
Thr Gln Pro Gly Ser Ile Thr Leu Arg Ile Thr Lys Asp Val Gly Tyr
            500                 505                 510
Gly Phe Ser Leu Phe Phe Thr Gly Ser Ser Ala Glu Leu Thr Asn Trp
        515                 520                 525
Gly Gly Gly Ile Glu Gly Thr Trp Thr Ser Gly His Val Trp Glu Val
    530                 535                 540
Thr Ile Pro Asp Pro Gly Asn Phe Glu Trp Lys Thr Arg Lys Gly Pro
545                 550                 555                 560
```

Thr Gly Gly Ser Gly Gln Asp Trp Glu Ser Asp Ser Asn His Asn Gln
565 570 575
Thr Asn Leu His Pro Ser Phe Asn Gly Gly Phe
580 585

<210> 611
<211> 3804
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 611
ttggaaacga cgatcattac attcttcttc aatgcgaaaa caaaagaggt gtacaccgac 60
tacaacccca atggctctga cggtgttgtt cagaaagacc gattaaaaca caatacatgg 120
gatacgctgt accgccagcc gtttggtgcg gtgaaagcgg gaacgaaggt gacgcttcgt 180
ctggcggcga aaaaaggcga tttgacgaag gcggatgtgt atgtgaaaaa tgcgacgacc 240
ggaaccgcca aagtctacac gatggaaaaa gtcggcgtgc tcggggatga tgaatattgg 300
gaggctgcgt ttacccctcc tgcgccgggg gtatacggtt ataaatttat cgctgtcgat 360
gccgggatga aggcggaata tggtgaagat acgaaagaag gacaatgggg gagggcggtc 420
gacaaaaatg cggggctgtt tcagctgacc gtctatgatc cgaattacaa aacaccagat 480
tggatgaaag aagcggttgt gtatcaaatt ttcccggatc ggttctttaa tggcaaccct 540
tcaaatgata acagcaagca gcaggcacgc ggggcgcagc cgattgagca tcgcgattgg 600
tcggatttgc ccgataatcc gcgcctgaaa gggacgagcg gctacgatgg cgacggtgaa 660
tggtcgaatg actttttcgg cggagacatc gccggaattg aacaaaagtt ggattatttg 720
cagtcgcttg gagtgaacac gatttactta aatccgatcg ccaatgcgcc atcgaaccat 780
aaatatgatg cgagcaatta caaagaattg gatccgatgt tcggttcccc ggaagaattc 840
caatcgtttg tgcaggcgct tgcgaaccgg gggatgcatc tcatcttaga cggggtgttc 900
aaccacgtat ccgacgattc gatttacttt gaccgctacc accgctatcc gaccgtcggt 960
gcgtatgaat attgggaagc ggtttacgat ttgatgaatg aaaaaggatt gagcgaggaa 1020
gaagcgcgga aacaagtgga agagaagttc aaacaagagg gacagacgtt cagcccgtat 1080
gggtttcacc tttggttcaa tattgaaaac aaaaaagtcg atggccatta tcaataccaa 1140
tcatggtggg gctatgacag tctgccggag tttaagtcgg tgacggggga aaaagtgccg 1200
catccgagtg aattgaacaa cgatgcgctc gcgaattaca ttttccgtga atcggattcg 1260
gtggcgaaaa gctggattgc cctcggcgcc tccggctggc ggttggatgt ggccaatgag 1320
gtggatccgg cgttttggcg cgagtttcgc caagaattgc ttcaagggtc gtacggtcgt 1380
ggtccgacgt taaaagaggg ggagcagccg ctcattttag gggaaatttg ggatgacgca 1440
tcgaaatatt ttctaggcga ccagtacgat tccgtgatga actaccggtt ccgcggggcg 1500
gtgcttgact ttttgaaaaa cggaaatgca gaagaggcgg acaagcggct gacggccata 1560
agggaagact acccaagtga agcgttttat gcgctgatga acttaatcgg ttcgcatgac 1620
acggcgcggg cggtctttct gcttgggaac ggaacggatt catccgagcg ggcggagctt 1680
gatccgaatt ataatgagga acttgggaaa aagcggctca agctggcggt gattttgcag 1740
atgggatacc cgggagcgcc gacgatttat tacggcgatg aagcgggagt aacaggctca 1800
agagacccag acaaccgccg cacgtatccg tggggcaaag aagatcaaaa tctgttgtcc 1860
cattatcaga aagtggggca cattcgccag caccatcaat cgttgttggc ccatggcgac 1920
atcaagacgg tgtatgcgca aggggatgta tacgtatttg cccgccaata cgggcgtgaa 1980
gtggcgctca ttgccctcaa ccgcggcaat gaggacaaga cggtggcgct tgacgtcgct 2040
tcgttgcttc cgaacggcac cgtgcttacg gatgagttgc atgatggcgg ggaagctacg 2100
gtcgctggcg gaacgttgac ggtcatgatt ccggccctgg atggacggat gatgtttggg 2160
acggtgacgg cggaaatgcc ggcagcagtc agcaatttgc aggcgagcgc ttcggatggc 2220
tgcgtgacgt taacgtggga aggaaatgcg tcgagatacc gaatttacga gtccacgtta 2280
aaaggtgccg gttatacgat ggtgcaagag acggaaacaa cttcggccac gatcggttcg 2340
ttgacgaacg gaacagccta ttactttgcc gttgcggcgg tcgatgaaaa cgggaatgaa 2400
tcaccgaagg tcgaaacgaa tcgcgtcgtt cctcattacc cgctgacgag cgacaatgtc 2460
cagttcgtga caacgttaag cgatgccaca ctggatttgt caaagccgca gcaagtggat 2520
gtccatgtca acatcgacaa tgtgacaagc aaaggagcag ctgatgggtt gcaagcggtg 2580
ttgcaagtga aaggcccgca tgacgaaaca tggaaagaat acagagcggc ttaccaagga 2640
caagacggcg acgccaacgt gttccgagct gccttcactc cgctcgccgc agggacgtat 2700
acgtatcgtt atgcgctgac gaccaacctc ggcgaggagt gggtgtatac agaagagaag 2760
caagtgacgt ttgcggcaga caacagcgac caaatagcgc cagcagacgc catcgagctg 2820
cggcagcctg cggttgaatc cggacaagtg aatttatcat ggacgtttgt tgggaaaaaa 2880
gatggggatg cttatttgtt agccatcgag cgcaacggtg atatcgtgca tacaaccact 2940
tcgatcggcg attcatttac agactacgat gtcgaaaacg gcaccgagta cacgtatgtt 3000
gtcaagttgt atgaccgcgc cggcaatgtt gtggcgtcaa acacggtcaa ggtgacgccg 3060
gacattgtga tggtgaaagt gatttttaaa gtgagggcgc cggattacac accgttggat 3120
gcccgaatta cgattccgaa cagcttgaac ggctggaaca caggggcctg ggagatgtcg 3180
cgcaacggtg cggtgacgcc cgattggcaa tttaccgtcg aggtgcagga aggggaaacg 3240

atcacctata agtatgtgaa aggcggatcg tgggatcaag aggggttggc cgaccatacg 3300
cgtgaggacg acaacgatga tgacgtgagc tactacggct atgggacgat tggcaccgac 3360
ttgaaagtga cggtccacaa tgaaggaaac aatacgatga ttgtgcaaga ccgcattttg 3420
cgctggatcg atatgccggt cgtcatcgaa gaggtgcaaa aacaaggaag tcaagtgacg 3480
atcaagggca atgccattaa aaacggtgtt ttgacgatca atggcgagcg ggtgccgatt 3540
gatggccgga tggcattctc gtacacgttt gcgccggcca gccatcaaaa agaagtgttg 3600
atccatatcg aaccatcggc cgaaagcaaa acagccattt tcaacaacga cggcggagcg 3660
attgcgaaaa acacaaaaga ttacgtgctg aatttagaaa cgaagcaatt caaaaagctt 3720
ctcgagagta cttctagagc ggccgcgggc ccatcgattt tccacccggg tggggtacca 3780
ggtaagtgta cccaattcgc ccta 3804

<210> 612
<211> 1268
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> DOMAIN
<222> (646)...(724)
<223> Alpha amylase, C-terminal all-beta domain

<221> DOMAIN
<222> (727)...(801)
<223> Fibronectin type III domain

<221> DOMAIN
<222> (169)...(629)
<223> Alpha amylase, catalytic domain

<400> 612
Met Glu Thr Thr Ile Ile Thr Phe Phe Phe Asn Ala Lys Thr Lys Glu
 1   5   10   15
Val Tyr Thr Asp Tyr Asn Pro Asn Gly Ser Asp Gly Val Val Gln Lys
            20   25   30
Asp Arg Leu Lys His Asn Thr Trp Asp Thr Leu Tyr Arg Gln Pro Phe
        35   40   45
Gly Ala Val Lys Ala Gly Thr Lys Val Thr Leu Arg Leu Ala Ala Lys
    50   55   60
Lys Gly Asp Leu Thr Lys Ala Asp Val Tyr Val Lys Asn Ala Thr Thr
65   70   75   80
Gly Thr Ala Lys Val Tyr Thr Met Glu Lys Val Gly Val Leu Gly Asp
                85   90   95
Asp Glu Tyr Trp Glu Ala Ala Phe Thr Pro Ser Ala Pro Gly Val Tyr
            100   105   110
Gly Tyr Lys Phe Ile Ala Val Asp Ala Gly Met Lys Ala Glu Tyr Gly
        115   120   125
Glu Asp Thr Lys Glu Gly Gln Trp Gly Arg Ala Val Asp Lys Asn Ala
    130   135   140
Gly Leu Phe Gln Leu Thr Val Tyr Asp Pro Asn Tyr Lys Thr Pro Asp
145   150   155   160
Trp Met Lys Glu Ala Val Val Tyr Gln Ile Phe Pro Asp Arg Phe Phe
                165   170   175
Asn Gly Asn Pro Ser Asn Asp Asn Ser Lys Gln Gln Ala Arg Gly Ala
            180   185   190
Gln Pro Ile Glu His Arg Asp Trp Ser Asp Leu Pro Asp Asn Pro Arg
        195   200   205
Leu Lys Gly Thr Ser Gly Tyr Asp Gly Asp Gly Glu Trp Ser Asn Asp
    210   215   220
Phe Phe Gly Gly Asp Ile Ala Gly Ile Glu Gln Lys Leu Asp Tyr Leu
225   230   235   240
Gln Ser Leu Gly Val Asn Thr Ile Tyr Leu Asn Pro Ile Ala Asn Ala
                245   250   255
Pro Ser Asn His Lys Tyr Asp Ala Ser Asn Tyr Lys Glu Leu Asp Pro
            260   265   270
Met Phe Gly Ser Pro Glu Glu Phe Gln Ser Phe Val Gln Ala Leu Ala
        275   280   285

```
Asn Arg Gly Met His Leu Ile Leu Asp Gly Val Phe Asn His Val Ser
    290                 295                 300
Asp Asp Ser Ile Tyr Phe Asp Arg Tyr His Arg Tyr Pro Thr Val Gly
305                 310                 315                 320
Ala Tyr Glu Tyr Trp Glu Ala Val Tyr Asp Leu Met Asn Glu Lys Gly
                325                 330                 335
Leu Ser Glu Glu Glu Ala Arg Lys Gln Val Glu Glu Lys Phe Lys Gln
            340                 345                 350
Glu Gly Gln Thr Phe Ser Pro Tyr Gly Phe His Leu Trp Phe Asn Ile
        355                 360                 365
Glu Asn Lys Lys Val Asp Gly His Tyr Gln Tyr Gln Ser Trp Trp Gly
    370                 375                 380
Tyr Asp Ser Leu Pro Glu Phe Lys Ser Val Thr Gly Glu Lys Val Pro
385                 390                 395                 400
His Pro Ser Glu Leu Asn Asn Asp Ala Leu Ala Asn Tyr Ile Phe Arg
                405                 410                 415
Glu Ser Asp Ser Val Ala Lys Ser Trp Ile Ala Leu Gly Ala Ser Gly
            420                 425                 430
Trp Arg Leu Asp Val Ala Asn Glu Val Asp Pro Ala Phe Trp Arg Glu
        435                 440                 445
Phe Arg Gln Glu Leu Leu Gln Gly Ser Tyr Gly Arg Gly Pro Thr Leu
    450                 455                 460
Lys Glu Gly Glu Gln Pro Leu Ile Leu Gly Glu Ile Trp Asp Asp Ala
465                 470                 475                 480
Ser Lys Tyr Phe Leu Gly Asp Gln Tyr Asp Ser Val Met Asn Tyr Arg
                485                 490                 495
Phe Arg Gly Ala Val Leu Asp Phe Leu Lys Asn Gly Asn Ala Glu Glu
            500                 505                 510
Ala Asp Lys Arg Leu Thr Ala Ile Arg Glu Asp Tyr Pro Ser Glu Ala
        515                 520                 525
Phe Tyr Ala Leu Met Asn Leu Ile Gly Ser His Asp Thr Ala Arg Ala
    530                 535                 540
Val Phe Leu Leu Gly Asn Gly Thr Asp Ser Ser Glu Arg Ala Glu Leu
545                 550                 555                 560
Asp Pro Asn Tyr Asn Glu Glu Leu Gly Lys Lys Arg Leu Lys Leu Ala
                565                 570                 575
Val Ile Leu Gln Met Gly Tyr Pro Gly Ala Pro Thr Ile Tyr Tyr Gly
            580                 585                 590
Asp Glu Ala Gly Val Thr Gly Ser Arg Asp Pro Asp Asn Arg Arg Thr
        595                 600                 605
Tyr Pro Trp Gly Lys Glu Asp Gln Asn Leu Leu Ser His Tyr Gln Lys
    610                 615                 620
Val Gly His Ile Arg Gln His His Gln Ser Leu Leu Ala His Gly Asp
625                 630                 635                 640
Ile Lys Thr Val Tyr Ala Gln Gly Asp Val Tyr Val Phe Ala Arg Gln
                645                 650                 655
Tyr Gly Arg Glu Val Ala Leu Ile Ala Leu Asn Arg Gly Asn Glu Asp
            660                 665                 670
Lys Thr Val Ala Leu Asp Val Ala Ser Leu Leu Pro Asn Gly Thr Val
        675                 680                 685
Leu Thr Asp Glu Leu His Asp Gly Gly Glu Ala Thr Val Ala Gly Gly
    690                 695                 700
Thr Leu Thr Val Met Ile Pro Ala Leu Asp Gly Arg Met Met Phe Gly
705                 710                 715                 720
Thr Val Thr Ala Glu Met Pro Ala Ala Val Ser Asn Leu Gln Ala Ser
                725                 730                 735
Ala Ser Asp Gly Cys Val Thr Leu Thr Trp Glu Gly Asn Ala Ser Arg
            740                 745                 750
Tyr Arg Ile Tyr Glu Ser Thr Leu Lys Gly Ala Gly Tyr Thr Met Val
        755                 760                 765
Gln Glu Thr Glu Thr Thr Ser Ala Thr Ile Gly Ser Leu Thr Asn Gly
    770                 775                 780
Thr Ala Tyr Tyr Phe Ala Val Ala Ala Val Asp Glu Asn Gly Asn Glu
785                 790                 795                 800
Ser Pro Lys Val Glu Thr Asn Arg Val Val Pro His Tyr Pro Leu Thr
                805                 810                 815
Ser Asp Asn Val Gln Phe Val Thr Thr Leu Ser Asp Ala Thr Leu Asp
            820                 825                 830
```

```
Leu Ser Lys Pro Gln Gln Val Asp Val His Val Asn Ile Asp Asn Val
        835                 840                 845
Thr Ser Lys Gly Ala Ala Asp Gly Leu Gln Ala Val Leu Gln Val Lys
    850                 855                 860
Gly Pro His Asp Glu Thr Trp Lys Glu Tyr Arg Ala Ala Tyr Gln Gly
865                 870                 875                 880
Gln Asp Gly Asp Ala Asn Val Phe Arg Ala Ala Phe Thr Pro Leu Ala
                885                 890                 895
Ala Gly Thr Tyr Thr Tyr Arg Tyr Ala Leu Thr Thr Asn Leu Gly Glu
            900                 905                 910
Glu Trp Val Tyr Thr Glu Glu Lys Gln Val Thr Phe Ala Ala Asp Asn
        915                 920                 925
Ser Asp Gln Ile Ala Pro Ala Asp Ala Ile Glu Leu Arg Gln Pro Ala
    930                 935                 940
Val Glu Ser Gly Gln Val Asn Leu Ser Trp Thr Phe Val Gly Lys Lys
945                 950                 955                 960
Asp Gly Asp Ala Tyr Leu Leu Ala Ile Glu Arg Asn Gly Asp Ile Val
                965                 970                 975
His Thr Thr Thr Ser Ile Gly Asp Ser Phe Thr Asp Tyr Asp Val Glu
            980                 985                 990
Asn Gly Thr Glu Tyr Thr Tyr Val Val Lys Leu Tyr Asp Arg Ala Gly
        995                 1000                1005
Asn Val Val Ala Ser Asn Thr Val Lys Val Thr Pro Asp Ile Val Met
    1010                1015                1020
Val Lys Val Ile Phe Lys Val Arg Ala Pro Asp Tyr Thr Pro Leu Asp
1025                1030                1035                1040
Ala Arg Ile Thr Ile Pro Asn Ser Leu Asn Gly Trp Asn Thr Gly Ala
                1045                1050                1055
Trp Glu Met Ser Arg Asn Gly Ala Val Thr Pro Asp Trp Gln Phe Thr
            1060                1065                1070
Val Glu Val Gln Glu Gly Glu Thr Ile Thr Tyr Lys Tyr Val Lys Gly
        1075                1080                1085
Gly Ser Trp Asp Gln Glu Gly Leu Ala Asp His Thr Arg Glu Asp Asp
    1090                1095                1100
Asn Asp Asp Asp Val Ser Tyr Tyr Gly Tyr Gly Thr Ile Gly Thr Asp
1105                1110                1115                1120
Leu Lys Val Thr Val His Asn Glu Gly Asn Asn Thr Met Ile Val Gln
                1125                1130                1135
Asp Arg Ile Leu Arg Trp Ile Asp Met Pro Val Val Ile Glu Glu Val
            1140                1145                1150
Gln Lys Gln Gly Ser Gln Val Thr Ile Lys Gly Asn Ala Ile Lys Asn
        1155                1160                1165
Gly Val Leu Thr Ile Asn Gly Glu Arg Val Pro Ile Asp Gly Arg Met
    1170                1175                1180
Ala Phe Ser Tyr Thr Phe Ala Pro Ala Ser His Gln Lys Glu Val Leu
1185                1190                1195                1200
Ile His Ile Glu Pro Ser Ala Glu Ser Lys Thr Ala Ile Phe Asn Asn
                1205                1210                1215
Asp Gly Gly Ala Ile Ala Lys Asn Thr Lys Asp Tyr Val Leu Asn Leu
            1220                1225                1230
Glu Thr Lys Gln Phe Lys Lys Leu Leu Glu Ser Thr Ser Arg Ala Ala
        1235                1240                1245
Ala Gly Pro Ser Ile Phe His Pro Gly Gly Val Pro Gly Lys Cys Thr
    1250                1255                1260
Gln Phe Ala Leu
1265

<210> 613
<211> 1764
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 613
atgcaaaaca cagcgaaaaa ctccatctgg cagagggtgc gccacagcgc cattgcctta      60
tccgctctca gcttatcctt tggcctgcag gccagcgagt taccacaaat tccaccacag     120
```

caggtgaaca acaccatgta ccaggcattt tattgggatg cctaccctgg cctttgggcc 180
aatttaccgg ccatggcggc ccctttggcc gagcgtggca ttacctcgat gtggttgccg 240
cccgccgcca aaggcatgaa tggtactttc agtgtcggtt acgatgtata cgatttctgg 300
gatctgggcg agtttaacca aaaaggcacc accgccaccc gttacggtac tcgtcagcag 360
ctgcaacaag cactgagtgc tctggaccaa ctgggtattc aggcctattt tgatgtggtg 420
tttaaccacc gcatgggcgc cgatgcacag gagaatattc ctggctttgg cctggcctgg 480
accgagtatc atctgcaagg tcgtcaggcg cattataccc agcaaaactg gggctacttg 540
tggcacgact ttgactggaa ctggaccgcg tttaatggct ccgacaatca gctctacccc 600
ggcaaatggt ggggcaatac cttccacttc ccttatttga tgggtgagga tgtcgattac 660
aaccgctttg aagtgcagca ggaaatgaaa gcctgggggg agtggatcat caacagcgtt 720
ggctttagcg gctttcggat ggatgccatt gcccatgttg ataccgattt tacccgtgac 780
tggatcaatc acgtacagtg ggccaccagc gaggatgtgt tctttgtcgc agaagcctgg 840
gtcagtgata taaacggcta tctggatgca gtcaatacgc cgcatttgcg cgcttttgat 900
ttcaatttgc gcgaagactt tgttgcttta agcagtggtg gcaaagacat gcgctggtgg 960
ggtggtctgg ttaatagcca gcaccgtgat cgggcggtca cttttgtcga taaccacgat 1020
accagccggg ccggcaaccc ttatggcatg ccgcaggtga tcaactacaa gaaccaggcc 1080
tacgcttaca ttctgttgcg tgagcatggg gtgccgactg tgtttgcccg cgattacgac 1140
gaatttggca tggcgccaac gctggataaa ttgattgaag cacggcgtta ctttgcctat 1200
ggtcctggcc atgagtactc cggcaatacc gaggccgtct acgcctatgt gcgcgaaggg 1260
cttagcactg tgccgggcac cggtctggtg atgctgatat cgggtcgaaa ctggggtggt 1320
cagcagtcgt tcaccatcaa cagccatcag ccgaatacca ccttttacga ttataccggc 1380
aatgtcagtg gcactgtgac caccaatgcg cagggttatg gcagcttccc ggtcactatg 1440
acggaaagta ccggttggtc agtctgggta ccacaatcca gtggtggcac tcagccggga 1500
tccattaccc tgcggataac caaggatgtt ggctatggtt tttcactgtt cttcactggc 1560
agcagtgcag agctcactaa ttggggcggc ggtattgaag gcacctggac cagtggccat 1620
gtctgggaag tgaccatccc ggatccgggc aactttgaat ggaaaacccg taaaggccca 1680
accggtggca gtggtcagga ctgggaaagt ggcagcaacc acaaccagac caatttgcac 1740
cccagtttta atggtgggtt ttaa 1764

<210> 614
<211> 587
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(31)

<221> DOMAIN
<222> (44)...(395)
<223> Alpha amylase, catalytic domain

<400> 614
Met Gln Asn Thr Ala Lys Asn Ser Ile Trp Gln Arg Val Arg His Ser
 1   5   10   15
Ala Ile Ala Leu Ser Ala Leu Ser Leu Ser Phe Gly Leu Gln Ala Ser
            20   25   30
Glu Leu Pro Gln Ile Pro Pro Gln Gln Val Asn Asn Thr Met Tyr Gln
        35   40   45
Ala Phe Tyr Trp Asp Ala Tyr Pro Gly Leu Trp Ala Asn Leu Pro Ala
    50   55   60
Met Ala Ala Pro Leu Ala Glu Arg Gly Ile Thr Ser Met Trp Leu Pro
65   70   75   80
Pro Ala Ala Lys Gly Met Asn Gly Thr Phe Ser Val Gly Tyr Asp Val
                85   90   95
Tyr Asp Phe Trp Asp Leu Gly Glu Phe Asn Gln Lys Gly Thr Thr Ala
            100   105   110
Thr Arg Tyr Gly Thr Arg Gln Gln Leu Gln Gln Ala Leu Ser Ala Leu
        115   120   125
Asp Gln Leu Gly Ile Gln Ala Tyr Phe Asp Val Val Phe Asn His Arg
    130   135   140
Met Gly Ala Asp Ala Gln Glu Asn Ile Pro Gly Phe Gly Leu Ala Trp
145   150   155   160
Thr Glu Tyr His Leu Gln Gly Arg Gln Ala His Tyr Thr Gln Gln Asn
                165   170   175
Trp Gly Tyr Leu Trp His Asp Phe Asp Trp Asn Trp Thr Ala Phe Asn

```
            180                 185                 190
Gly Ser Asp Asn Gln Leu Tyr Pro Gly Lys Trp Trp Gly Asn Thr Phe
        195                 200                 205
His Phe Pro Tyr Leu Met Gly Glu Asp Val Asp Tyr Asn Arg Phe Glu
    210                 215                 220
Val Gln Gln Glu Met Lys Ala Trp Gly Glu Trp Ile Ile Asn Ser Val
225                 230                 235                 240
Gly Phe Ser Gly Phe Arg Met Asp Ala Ile Ala His Val Asp Thr Asp
                245                 250                 255
Phe Thr Arg Asp Trp Ile Asn His Val Gln Trp Ala Thr Ser Glu Asp
            260                 265                 270
Val Phe Phe Val Ala Glu Ala Trp Val Ser Asp Ile Asn Gly Tyr Leu
        275                 280                 285
Asp Ala Val Asn Thr Pro His Leu Arg Ala Phe Asp Phe Asn Leu Arg
    290                 295                 300
Glu Asp Phe Val Ala Leu Ser Ser Gly Gly Lys Asp Met Arg Trp Trp
305                 310                 315                 320
Gly Gly Leu Val Asn Ser Gln His Arg Asp Arg Ala Val Thr Phe Val
                325                 330                 335
Asp Asn His Asp Thr Ser Arg Ala Gly Asn Pro Tyr Gly Met Pro Gln
            340                 345                 350
Val Ile Asn Tyr Lys Asn Gln Ala Tyr Ala Tyr Ile Leu Leu Arg Glu
        355                 360                 365
His Gly Val Pro Thr Val Phe Ala Arg Asp Tyr Asp Glu Phe Gly Met
    370                 375                 380
Ala Pro Thr Leu Asp Lys Leu Ile Glu Ala Arg Arg Tyr Phe Ala Tyr
385                 390                 395                 400
Gly Pro Gly His Glu Tyr Ser Gly Asn Thr Glu Ala Val Tyr Ala Tyr
                405                 410                 415
Val Arg Glu Gly Leu Ser Thr Val Pro Gly Thr Gly Leu Val Met Leu
            420                 425                 430
Ile Ser Gly Arg Asn Trp Gly Gly Gln Gln Ser Phe Thr Ile Asn Ser
        435                 440                 445
His Gln Pro Asn Thr Thr Phe Tyr Asp Tyr Thr Gly Asn Val Ser Gly
    450                 455                 460
Thr Val Thr Thr Asn Ala Gln Gly Tyr Gly Ser Phe Pro Val Thr Met
465                 470                 475                 480
Thr Glu Ser Thr Gly Trp Ser Val Trp Val Pro Gln Ser Ser Gly Gly
                485                 490                 495
Thr Gln Pro Gly Ser Ile Thr Leu Arg Ile Thr Lys Asp Val Gly Tyr
            500                 505                 510
Gly Phe Ser Leu Phe Phe Thr Gly Ser Ser Ala Glu Leu Thr Asn Trp
        515                 520                 525
Gly Gly Gly Ile Glu Gly Thr Trp Thr Ser Gly His Val Trp Glu Val
    530                 535                 540
Thr Ile Pro Asp Pro Gly Asn Phe Glu Trp Lys Thr Arg Lys Gly Pro
545                 550                 555                 560
Thr Gly Gly Ser Gly Gln Asp Trp Glu Ser Gly Ser Asn His Asn Gln
                565                 570                 575
Thr Asn Leu His Pro Ser Phe Asn Gly Gly Phe
            580                 585
```

<210> 615
<211> 1752
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 615
atgtccgaaa gaggcgtcag gagagccgtg cgcaccgccc tggtcgggtt ggcggcggcc 60
gccaccgccg cggtgacgct cggcgctccc accgcccagg ccgcgccggc cgggaacaag 120
gacgtcaccg ccgtcctgtt ctcctggaac ttcaactcca tcgcccgcga gtgcaccgag 180
cgactcggcc ccgcgggcta cggcttcgtg caggtctcgc cgccccagga gcacatccag 240
ggctccgcgt ggtggaccca gtaccagccc gtgagttaca acatctccgg ccggctcggc 300
aacgagcagc agttccgctc gatggtcaac acctgtaaca acgcgggtgt cggcgtcatc 360
gtggacgcgg tcatcaacca catgtccgcc ggctcgggca ccggcaccgg cggcacctcc 420

tacagcaagt acaactaccc cggcctctac agcggcaacg acttcggttc ctgccgcgag 480
gacatccggg gcagcgacta caccaacgac cgctggcggg tccagaactg cgagctggtc 540
ggcctcgccg acctggacac cggcagcgcc tacgtccagc agcagatcgc cggctacatg 600
aaccgtctgc tgggctgggg cgtcgccggc ttccgggtgg acgccgtcaa gcacatcccg 660
gcccggcaca tggagcagat ccgcgcccgg gtcaacggcg gcgacgtctt ctggaaccag 720
gaggtcatct acggcgccgg cgaggcgatc acgcccgagg agtacctcaa caccggtcac 780
gtgcaggagt tccgctacgc gttcgacctc aagcggatgt tccagggcga ccggatcgcc 840
aatctgcaga acttcggcga gtcatggggc tacatgccct ccaaccgctc cggcgtcttc 900
gtcgacaacc acgacaccga gcgcaacggc tccaccctca gctacaagga caacgccgcc 960
tacaccctgg ccaacgtctt catgctggcc tggccctacg gcagcccgga cgtgcactcc 1020
ggctacgagt tcaccaactt cgacgccggt ccgcccaaca acggcaacgt gaccgcctgc 1080
caccagagcg ggtggaagtg ccagcacgcc tggcaggaga tctcctccat ggtgggcttc 1140
cgcaacgcca cccgcggcca ggccgtcacc aactggtggt ccaacggcaa caacgccatc 1200
gccttcggcc ggggcaaccg cggctacgtg gcgatcaacc acgagaacac cacactgaac 1260
cggaccttcc agacctccct gccggccggt gactactgca acgtgcagaa cggcaccacc 1320
gtcaccgtca acggcgccgg gcagttcacc gcgagcctag gcgcccgcac cgccctcgcc 1380
ctgcacgtcg acgcccgcaa ttgcgccggc ggcggcaccg ggggcaacgg cggcaccggg 1440
ggaggcgaga acccggtgac cggcggtgcc tccttcggcg tcgacgcgac cacgcagatg 1500
ggccagaaca tccacgtcgc gggcaacatc cccgcgctgg gcgactggaa caccgccaac 1560
gccccccgca tgagcgccga cacctacccg gtctggcggc tggaactgaa cctgccgccc 1620
ggcaccacct tccagtacaa gtacatccgc agggacgcca acggcaacgt gacctgggaa 1680
tccgggaaca accgcaccgc caccgtgccc tcctccggcc gggtgacgct caacgacacc 1740
tggcgcaact ga 1752

<210> 616
<211> 583
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(34)

<221> DOMAIN
<222> (390)...(466)
<223> Alpha amylase, C-terminal all-beta domain

<221> DOMAIN
<222> (487)...(579)
<223> Starch binding domain

<221> DOMAIN
<222> (41)...(381)
<223> Alpha amylase, catalytic domain

<400> 616
Met Ser Glu Arg Gly Val Arg Arg Ala Val Arg Thr Ala Leu Val Gly
 1   5   10   15
Leu Ala Ala Ala Ala Thr Ala Ala Val Thr Leu Gly Ala Pro Thr Ala
            20   25   30
Gln Ala Ala Pro Ala Gly Asn Lys Asp Val Thr Ala Val Leu Phe Ser
        35   40   45
Trp Asn Phe Asn Ser Ile Ala Arg Glu Cys Thr Glu Arg Leu Gly Pro
    50   55   60
Ala Gly Tyr Gly Phe Val Gln Val Ser Pro Pro Gln Glu His Ile Gln
65   70   75   80
Gly Ser Ala Trp Thr Gln Tyr Gln Pro Val Ser Tyr Asn Ile Ser
                85   90   95
Gly Arg Leu Gly Asn Glu Gln Gln Phe Arg Ser Met Val Asn Thr Cys
            100   105   110
Asn Asn Ala Gly Val Gly Val Ile Val Asp Ala Val Ile Asn His Met
        115   120   125
Ser Ala Gly Ser Gly Thr Gly Thr Gly Gly Thr Ser Tyr Ser Lys Tyr
    130   135   140
Asn Tyr Pro Gly Leu Tyr Ser Gly Asn Asp Phe Gly Ser Cys Arg Glu
145   150   155   160

```
Asp Ile Arg Gly Ser Asp Tyr Thr Asn Asp Arg Trp Arg Val Gln Asn
                165                 170                 175
Cys Glu Leu Val Gly Leu Ala Asp Leu Asp Thr Gly Ser Ala Tyr Val
            180                 185                 190
Gln Gln Gln Ile Ala Gly Tyr Met Asn Arg Leu Leu Gly Trp Gly Val
        195                 200                 205
Ala Gly Phe Arg Val Asp Ala Val Lys His Ile Pro Ala Arg His Met
    210                 215                 220
Glu Gln Ile Arg Ala Arg Val Asn Gly Gly Asp Val Phe Trp Asn Gln
225                 230                 235                 240
Glu Val Ile Tyr Gly Ala Gly Glu Ala Ile Thr Pro Glu Glu Tyr Leu
                245                 250                 255
Asn Thr Gly His Val Gln Glu Phe Arg Tyr Ala Phe Asp Leu Lys Arg
            260                 265                 270
Met Phe Gln Gly Asp Arg Ile Ala Asn Leu Gln Asn Phe Gly Glu Ser
        275                 280                 285
Trp Gly Tyr Met Pro Ser Asn Arg Ser Gly Val Phe Val Asp Asn His
    290                 295                 300
Asp Thr Glu Arg Asn Gly Ser Thr Leu Ser Tyr Lys Asp Asn Ala Ala
305                 310                 315                 320
Tyr Thr Leu Ala Asn Val Phe Met Leu Ala Trp Pro Tyr Gly Ser Pro
                325                 330                 335
Asp Val His Ser Gly Tyr Glu Phe Thr Asn Phe Asp Ala Gly Pro Pro
            340                 345                 350
Asn Asn Gly Asn Val Thr Ala Cys His Gln Ser Gly Trp Lys Cys Gln
        355                 360                 365
His Ala Trp Gln Glu Ile Ser Ser Met Val Gly Phe Arg Asn Ala Thr
    370                 375                 380
Arg Gly Gln Ala Val Thr Asn Trp Trp Ser Asn Gly Asn Asn Ala Ile
385                 390                 395                 400
Ala Phe Gly Arg Gly Asn Arg Gly Tyr Val Ala Ile Asn His Glu Asn
                405                 410                 415
Thr Thr Leu Asn Arg Thr Phe Gln Thr Ser Leu Pro Ala Gly Asp Tyr
            420                 425                 430
Cys Asn Val Gln Asn Gly Thr Thr Val Thr Val Asn Gly Ala Gly Gln
        435                 440                 445
Phe Thr Ala Ser Leu Gly Ala Arg Thr Ala Leu Ala Leu His Val Asp
    450                 455                 460
Ala Arg Asn Cys Ala Gly Gly Gly Thr Gly Gly Asn Gly Gly Thr Gly
465                 470                 475                 480
Gly Gly Glu Asn Pro Val Thr Gly Gly Ala Ser Phe Gly Val Asp Ala
                485                 490                 495
Thr Thr Gln Met Gly Gln Asn Ile His Val Ala Gly Asn Ile Pro Ala
            500                 505                 510
Leu Gly Asp Trp Asn Thr Ala Asn Ala Pro Arg Met Ser Ala Asp Thr
        515                 520                 525
Tyr Pro Val Trp Arg Leu Glu Leu Asn Leu Pro Pro Gly Thr Thr Phe
    530                 535                 540
Gln Tyr Lys Tyr Ile Arg Arg Asp Ala Asn Gly Asn Val Thr Trp Glu
545                 550                 555                 560
Ser Gly Asn Asn Arg Thr Ala Thr Val Pro Ser Ser Gly Arg Val Thr
                565                 570                 575
Leu Asn Asp Thr Trp Arg Asn
            580

<210> 617
<211> 1668
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 617
atgaaccgat acctgcgact ggccgcgttg acgctggccc tggcgccgct ggcctacccc      60
tggggtaatc tggcccgcgc cgccgatgca ccgggcaaga ccgccagcgg cgtgcgctac     120
cacggcggcg acgaaatcat cctccagggc tttcactgga acaccgtgcg tacctcgagc     180
aactggtacg cgacgctggc cagcatggcg ccgaccctgg ccgccgatgg tttcagcgca     240
```

```
atctggatgc cggtgccctg gcgcgacttc tccagctgga gcgaccccgg caacggcacc     300
tcgggcggcg gcgaaggcta cttctggcac gacttcaaca agaacggccg ctatggcagc     360
gacagcctgc tcaaacaggc cgccggcgcg ctcaatgcgg ccggggtcaa acccatctac     420
gacgtggtgc ccaaccacat gaaccgcggc tacccggaca aggaaatcaa cctgccggcc     480
ggccaggggc tatggcgcca cgactgcaat gacccgggca actacgccaa cgactgcgac     540
gacggcgacc gcttcatggg cggcgacgcc gacctcaata ccggccaccc gcagaactac     600
gcgatgttcc gcgacgagtt cgccaggctg cgcagccagt atggcgccgg gggctttcgc     660
ttcgacttcg tccgcggcta tgccggcgag cgggtcgcca actggatgag cgatgcccac     720
gacaacggct tctgcctcgg cgagctgtgg aaggcgccgg gcgaataccc aagctgggac     780
tggcgcaacg gcgccagctg gcagcagatc ctcaaggact ggtccgatcg cgccaagtgc     840
acggtgttcg acttcgccct caaggagcgc atgcagaacg gcggcatcgc cgactggcgc     900
cacggcctca acggcaaccc cgatgcgcgc tggcgcgagg tggcggtgac cttcgtcgac     960
aaccacgaca ccggctattc gcccggcccg cacggcggtc agcatcactg gccgctgccg    1020
gacgcacggc tgaagcaagc ctatgcctac atcctcagca gcccgggcac gccggtggtg    1080
tactggccgc acatgtacga ctggggccat ggcgacttca tccgccagtt gatccagatc    1140
cgccgtgcgg cgggggtcaa ggcggcctcg gcgatccagt tccactcggg tttctccgga    1200
ctggtggcga cagtcagcgg cagccagcag caactgctga tcgccctgga ctccaacctg    1260
tccacgccca gccaggtggc cagcggcgac ttcacccagg cgctcaacac tgacaacggc    1320
gcgatccgta tctggcgcag cggtcagagt gacgatgagg agcagggcga tctggtcagc    1380
gtcaacttcc gctgtgacaa cggcgtgacc cagtggggcg acagcgtcta cgcggtgggc    1440
aacgtcgccc aactgggcaa ctggagcccg gccagcgccg tgcgcctgac cgataccagc    1500
gcctacccga cctggaaagg cagcatcgcc ctgccggccg gccagcaggt gcagtggaaa    1560
tgcatcgtgc gcagcgaaag caacccgacc caggtcaaga cctggcagcc gggcggcaac    1620
aacgccgtga cggtgacggc aggggccagc accgcgggga gtttctaa                1668

<210> 618
<211> 555
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(27)

<221> DOMAIN
<222> (458)...(554)
<223> Starch binding domain

<221> DOMAIN
<222> (45)...(395)
<223> Alpha amylase, catalytic domain

<400> 618
Met Asn Arg Tyr Leu Arg Leu Ala Ala Leu Thr Leu Ala Leu Ala Pro
 1               5                  10                  15
Leu Ala Tyr Pro Trp Gly Asn Leu Ala Arg Ala Ala Asp Ala Pro Gly
            20                  25                  30
Lys Thr Ala Ser Gly Val Arg Tyr His Gly Gly Asp Glu Ile Ile Leu
        35                  40                  45
Gln Gly Phe His Trp Asn Thr Val Arg Thr Ser Ser Asn Trp Tyr Ala
    50                  55                  60
Thr Leu Ala Ser Met Ala Pro Thr Leu Ala Ala Asp Gly Phe Ser Ala
65                  70                  75                  80
Ile Trp Met Pro Val Pro Trp Arg Asp Phe Ser Ser Trp Ser Asp Pro
                85                  90                  95
Gly Asn Gly Thr Ser Gly Gly Gly Glu Gly Tyr Phe Trp His Asp Phe
            100                 105                 110
Asn Lys Asn Gly Arg Tyr Gly Ser Asp Ser Leu Leu Lys Gln Ala Ala
        115                 120                 125
Gly Ala Leu Asn Ala Ala Gly Val Lys Pro Ile Tyr Asp Val Val Pro
    130                 135                 140
Asn His Met Asn Arg Gly Tyr Pro Asp Lys Glu Ile Asn Leu Pro Ala
145                 150                 155                 160
Gly Gln Gly Leu Trp Arg His Asp Cys Asn Asp Pro Gly Asn Tyr Ala
                165                 170                 175
Asn Asp Cys Asp Asp Gly Asp Arg Phe Met Gly Gly Asp Ala Asp Leu
```

```
            180                 185                 190
Asn Thr Gly His Pro Gln Asn Tyr Ala Met Phe Arg Asp Glu Phe Ala
        195                 200                 205
Arg Leu Arg Ser Gln Tyr Gly Ala Gly Gly Phe Arg Phe Asp Phe Val
    210                 215                 220
Arg Gly Tyr Ala Gly Glu Arg Val Ala Asn Trp Met Ser Asp Ala His
225                 230                 235                 240
Asp Asn Gly Phe Cys Leu Gly Glu Leu Trp Lys Ala Pro Gly Glu Tyr
                245                 250                 255
Pro Ser Trp Asp Trp Arg Asn Gly Ala Ser Trp Gln Gln Ile Leu Lys
            260                 265                 270
Asp Trp Ser Asp Arg Ala Lys Cys Thr Val Phe Asp Phe Ala Leu Lys
        275                 280                 285
Glu Arg Met Gln Asn Gly Gly Ile Ala Asp Trp Arg His Gly Leu Asn
    290                 295                 300
Gly Asn Pro Asp Ala Arg Trp Arg Glu Val Ala Val Thr Phe Val Asp
305                 310                 315                 320
Asn His Asp Thr Gly Tyr Ser Pro Gly Pro His Gly Gly Gln His His
                325                 330                 335
Trp Pro Leu Pro Asp Ala Arg Leu Lys Gln Ala Tyr Ala Tyr Ile Leu
            340                 345                 350
Ser Ser Pro Gly Thr Pro Val Val Tyr Trp Pro His Met Tyr Asp Trp
        355                 360                 365
Gly His Gly Asp Phe Ile Arg Gln Leu Ile Gln Ile Arg Arg Ala Ala
    370                 375                 380
Gly Val Lys Ala Ala Ser Ala Ile Gln Phe His Ser Gly Phe Ser Gly
385                 390                 395                 400
Leu Val Ala Thr Val Ser Gly Ser Gln Gln Gln Leu Leu Ile Ala Leu
                405                 410                 415
Asp Ser Asn Leu Ser Thr Pro Ser Gln Val Ala Ser Gly Asp Phe Thr
            420                 425                 430
Gln Ala Leu Asn Thr Asp Asn Gly Ala Ile Arg Ile Trp Arg Ser Gly
        435                 440                 445
Gln Ser Asp Asp Glu Glu Gln Gly Asp Leu Val Ser Val Asn Phe Arg
    450                 455                 460
Cys Asp Asn Gly Val Thr Gln Trp Gly Asp Ser Val Tyr Ala Val Gly
465                 470                 475                 480
Asn Val Ala Gln Leu Gly Asn Trp Ser Pro Ala Ser Ala Val Arg Leu
                485                 490                 495
Thr Asp Thr Ser Ala Tyr Pro Thr Trp Lys Gly Ser Ile Ala Leu Pro
            500                 505                 510
Ala Gly Gln Gln Val Gln Trp Lys Cys Ile Val Arg Ser Glu Ser Asn
        515                 520                 525
Pro Thr Gln Val Lys Thr Trp Gln Pro Gly Gly Asn Asn Ala Val Thr
    530                 535                 540
Val Thr Ala Gly Ala Ser Thr Ala Gly Ser Phe
545                 550                 555

<210> 619
<211> 1716
<212> DNA
<213> Bacteria

<400> 619
atggcaagga agtccgtggc cgccgcactc gccctcgtgg cgggagccgc cgccgtggcc      60
gtcacgggaa acacggcggc gcaggccaca ccgcccgggg agaaggacgt caccgcggtg     120
atgttcgaat ggaacttcgc ctccgtcgcc cgcgagtgca ccgaccgcct cggccccgcc     180
ggatacggat acgtccaggt ctccccgccc caggaacacc tccagggcgg ccagtggtgg     240
acctcctacc agccggtcag ctacaagatc gccggacgcc tcggtgaccg caccgccttc     300
aagaagatga tcgacacctg ccacgcggcc ggcgtcaagg tcgtcgccga ctccgtcatc     360
aaccacatgg cgaacggctc cggcacggga acgggcggga ccccgttctc caagtacgac     420
tacccgggcc tctactccgg cgccgacatg gacgactgcc gcgcgacgat ctccacctac     480
caggaccggg gcaacgtcca gaactgcgaa ctggtccagc tccccgacct cgacaccggc     540
gaggaccacg tccgcgggaa gatagccggc tacctcaacg acctgctctc cctcggcgtc     600
gacggcttcc gcgtcgacgc cgccaagcac atgccggccg ccgacctcgc cgccatcaag     660
gcgcggctca gcaaccccgg cgcgtactgg aagcaggagg ccatcttcgg cgcgggcgag     720
gccgtctccc cgagcgagta cctcggcaac ggcgacgtcc aggagttccg ctacgcccgc     780
gacctcaagc gcgtcctgca gaacgagaag ctcgcctacc tcaagaactt cggcgaggcc     840
```

```
tgggggtaca tgccctccgc ccagtccggc gtcttcgtcg acaaccacga caccgagcgc     900
ggcggcgaca ccctcagcta caaggacggc gcgaactaca ccctcgcctc cgtcttcatg     960
ctcgcctggc cctacggctc cccggacgtc cactccggct acgagtggac cgacaaggac    1020
gccggcccgc cgagcggcgg acaggtcaac gcctgctaca ccgacggctg gaagtgccag    1080
cacgcctggc gcgagatctc ctccatggtc gccttccgca acaccgcgcg cggccaggcc    1140
gtcaccggct ggtgggacaa cggcaacaac gcgatcgctt tcggccgcgg ctccaaggcg    1200
tacgtggcca tcaaccacga gacctcggcg ctcacccgca cctaccagac ctcgctgccc    1260
gccggcacct actgcgacgt ccagtcgaac acccccgtga cggtgaacgg ctccggccag    1320
ttcaccgcca ccctcggcgc ccacaccgcc ctcgccctgc acgtcggcgc gatgaactgc    1380
ggctccggca cgaccacgcc accgccgacg accccgcccg ccaccccggg cgcgtccttc    1440
aacgccaccg ccaccaccgt ggtcggtcag gacatctacg tcaccggcaa ccgcgccgag    1500
ctcggcaact ggtccccggc cgccgccctc aagctcgacc ccgcgacgta ccccgtctgg    1560
aagctgaccg tcggcctgcc cgccggaacc gccttcgagt acaagtacct ccgcaaggac    1620
gccgccggaa acgtcacctg ggagtccggc gccaaccgca ccgcgaccgt ccccgcctcc    1680
gggcgcatcg tcctgaacga caccttccgc agctga                              1716

<210> 620
<211> 571
<212> PRT
<213> Bacteria

<220>
<221> SIGNAL
<222> (1)...(29)

<221> DOMAIN
<222> (475)...(567)
<223> Starch binding domain

<221> DOMAIN
<222> (382)...(458)
<223> Alpha amylase, C-terminal all-beta domain

<221> DOMAIN
<222> (36)...(373)
<223> Alpha amylase, catalytic domain

<400> 620
Met Ala Arg Lys Ser Val Ala Ala Ala Leu Ala Leu Val Ala Gly Ala
 1               5                  10                  15
Ala Ala Val Ala Val Thr Gly Asn Thr Ala Ala Gln Ala Thr Pro Pro
            20                  25                  30
Gly Glu Lys Asp Val Thr Ala Val Met Phe Glu Trp Asn Phe Ala Ser
        35                  40                  45
Val Ala Arg Glu Cys Thr Asp Arg Leu Gly Pro Ala Gly Tyr Gly Tyr
    50                  55                  60
Val Gln Val Ser Pro Pro Gln Glu His Leu Gln Gly Gly Gln Trp Trp
65                  70                  75                  80
Thr Ser Tyr Gln Pro Val Ser Tyr Lys Ile Ala Gly Arg Leu Gly Asp
                85                  90                  95
Arg Thr Ala Phe Lys Lys Met Ile Asp Thr Cys His Ala Ala Gly Val
            100                 105                 110
Lys Val Val Ala Asp Ser Val Ile Asn His Met Ala Asn Gly Ser Gly
        115                 120                 125
Thr Gly Thr Gly Gly Thr Pro Phe Ser Lys Tyr Asp Tyr Pro Gly Leu
    130                 135                 140
Tyr Ser Gly Ala Asp Met Asp Asp Cys Arg Ala Thr Ile Ser Thr Tyr
145                 150                 155                 160
Gln Asp Arg Gly Asn Val Gln Asn Cys Glu Leu Val Gln Leu Pro Asp
                165                 170                 175
Leu Asp Thr Gly Glu Asp His Val Arg Gly Lys Ile Ala Gly Tyr Leu
            180                 185                 190
Asn Asp Leu Leu Ser Leu Gly Val Asp Gly Phe Arg Val Asp Ala Ala
        195                 200                 205
Lys His Met Pro Ala Ala Asp Leu Ala Ala Ile Lys Ala Arg Leu Ser
    210                 215                 220
Asn Pro Gly Ala Tyr Trp Lys Gln Glu Ala Ile Phe Gly Ala Gly Glu
225                 230                 235                 240
```

```
Ala Val Ser Pro Ser Glu Tyr Leu Gly Asn Gly Asp Val Gln Glu Phe
                245                 250                 255
Arg Tyr Ala Arg Asp Leu Lys Arg Val Leu Gln Asn Glu Lys Leu Ala
            260                 265                 270
Tyr Leu Lys Asn Phe Gly Glu Ala Trp Gly Tyr Met Pro Ser Ala Gln
        275                 280                 285
Ser Gly Val Phe Val Asp Asn His Asp Thr Glu Arg Gly Gly Asp Thr
    290                 295                 300
Leu Ser Tyr Lys Asp Gly Ala Asn Tyr Thr Leu Ala Ser Val Phe Met
305                 310                 315                 320
Leu Ala Trp Pro Tyr Gly Ser Pro Asp Val His Ser Gly Tyr Glu Trp
                325                 330                 335
Thr Asp Lys Asp Ala Gly Pro Pro Ser Gly Gly Gln Val Asn Ala Cys
            340                 345                 350
Tyr Thr Asp Gly Trp Lys Cys Gln His Ala Trp Arg Glu Ile Ser Ser
        355                 360                 365
Met Val Ala Phe Arg Asn Thr Ala Arg Gly Gln Ala Val Thr Gly Trp
    370                 375                 380
Trp Asp Asn Gly Asn Asn Ala Ile Ala Phe Gly Arg Gly Ser Lys Ala
385                 390                 395                 400
Tyr Val Ala Ile Asn His Glu Thr Ser Ala Leu Thr Arg Thr Tyr Gln
                405                 410                 415
Thr Ser Leu Pro Ala Gly Thr Tyr Cys Asp Val Gln Ser Asn Thr Pro
            420                 425                 430
Val Thr Val Asn Gly Ser Gly Gln Phe Thr Ala Thr Leu Gly Ala His
        435                 440                 445
Thr Ala Leu Ala Leu His Val Gly Ala Met Asn Cys Gly Ser Gly Thr
    450                 455                 460
Thr Thr Pro Pro Pro Thr Thr Pro Pro Ala Thr Pro Gly Ala Ser Phe
465                 470                 475                 480
Asn Ala Thr Ala Thr Thr Val Val Gly Gln Asp Ile Tyr Val Thr Gly
                485                 490                 495
Asn Arg Ala Glu Leu Gly Asn Trp Ser Pro Ala Ala Ala Leu Lys Leu
            500                 505                 510
Asp Pro Ala Thr Tyr Pro Val Trp Lys Leu Thr Val Gly Leu Pro Ala
        515                 520                 525
Gly Thr Ala Phe Glu Tyr Lys Tyr Leu Arg Lys Asp Ala Ala Gly Asn
    530                 535                 540
Val Thr Trp Glu Ser Gly Ala Asn Arg Thr Ala Thr Val Pro Ala Ser
545                 550                 555                 560
Gly Arg Ile Val Leu Asn Asp Thr Phe Arg Ser
                565                 570
```

<210> 621
<211> 1770
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

<400> 621

```
atgcaaaaca cagcgaaaaa ctccatctgg cagagggtgc gccacagcgc cattgcctta      60
tccgctctca gcttatcctt tggcctgcag gccagcgagt taccacaaat tccaccacag     120
caggtgaaca acaccatgta ccaggcattt tattgggatg cctaccctgg cctttgggcc     180
aatttaccgg ccatggcggc ccctttggcc gaacggggca ttacctcgat gtggttgccg     240
cccgccgcca aaggcatgaa tggtactttc agtgtcggtt acgatgtata cgatttctgg     300
gatctgggcg agtttaacca aaaaggcacc accgccaccc gttacggtac tcgtcagcag     360
ctgcaacaag cactgagtgc tctggaccaa ctgggtattc aggcctattt tgatgtggtg     420
tttaaccacc gcatgggcgc cgatgcacag gagaatattc ctggctttgg cctggcctgg     480
accgagtatc atctgcaagg tcgtcaggcg cattataccc agcaaaactg gggctacttg     540
tggcacgact ttgactggaa ctggaccgcg tttaatggct ccgacaatca gctctacccc     600
ggcaaatggt ggggcaatac cttccacttc ccttatttga tgggtgagga tgtcgattac     660
aaccgctttg aagtgcagca ggaaatgaaa gcctggggcg agtggatcat caacagcgtt     720
ggctttagcg gctttcggat ggacgccatc gcacacgtgg ataccgattt tacccgtgac     780
tggatcaatc acgtacagtg ggccaccagc gaggatgtat tctttgtcgc agaagcctgg     840
gtcagtgata tcaacggcta tctggatgca gtcaatacgc cgcatttgcg cgcttttgat     900
ttcaatttgc gcgaagactt tgttgcttta agcagtggtg gcaaagacat gcgctggtgg     960
```

```
ggtggtctgg ttaatagcca gcaccgtgat cgggcggtca cttttgtcga taaccacgat    1020
accagccggg ccggcaaccc ttatggcatg ccgcaggtga tcaactacaa gaaccaggcc    1080
tacgcttaca ttctgttgcg tgagcatggg gtgccgactg tgtttgcccg cgattacgac    1140
gaatttggca tggcgccaac gctggataaa ttgattgagg cgcgccgcta ctttgcctat    1200
ggtcctggtc atgagtactc cggcaatacc gaggccgttt acgcctatgt gcgcgagggg    1260
cttagcactg tgccgggcac cggtctggtg atgctgatat cgggtcgaaa ctggggtggt    1320
cagcagtcgt tcaccatcaa cagccatcag ccgaatacca ccttttacga ttataccggc    1380
aatgtcagcg gcacggtgac caccaatgcg cagggttatg gcagcttccc ggtcactatg    1440
acggaaagta ccggttggtc agtctgggta ccacaatcca gtggtggcac tcagccggga    1500
tccattaccc tgcggatgac caaggatgtt ggctatggtt tttcactgtt cttcactggc    1560
agcagtgcag agctcaccaa ttggggcggc ggtattgaag gcacctggac cagtggccat    1620
gtctgggaag tgaccatccc ggatccgggc aactttgaat ggaaaacccg taaaggccca    1680
accggcggca gtggtcagga ctgggaaagt ggtagcaacc acaatcagac caatttgcac    1740
cccagtttta atggtgggtt tagatctcgc                                     1770

<210> 622
<211> 590
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(31)

<221> DOMAIN
<222> (44)...(395)
<223> Alpha amylase, catalytic domain

<400> 622
Met Gln Asn Thr Ala Lys Asn Ser Ile Trp Gln Arg Val Arg His Ser
 1               5                  10                  15
Ala Ile Ala Leu Ser Ala Leu Ser Leu Ser Phe Gly Leu Gln Ala Ser
            20                  25                  30
Glu Leu Pro Gln Ile Pro Pro Gln Gln Val Asn Asn Thr Met Tyr Gln
        35                  40                  45
Ala Phe Tyr Trp Asp Ala Tyr Pro Gly Leu Trp Ala Asn Leu Pro Ala
    50                  55                  60
Met Ala Ala Pro Leu Ala Glu Arg Gly Ile Thr Ser Met Trp Leu Pro
65                  70                  75                  80
Pro Ala Ala Lys Gly Met Asn Gly Thr Phe Ser Val Gly Tyr Asp Val
                85                  90                  95
Tyr Asp Phe Trp Asp Leu Gly Glu Phe Asn Gln Lys Gly Thr Thr Ala
            100                 105                 110
Thr Arg Tyr Gly Thr Arg Gln Gln Leu Gln Gln Ala Leu Ser Ala Leu
        115                 120                 125
Asp Gln Leu Gly Ile Gln Ala Tyr Phe Asp Val Val Phe Asn His Arg
    130                 135                 140
Met Gly Ala Asp Ala Gln Glu Asn Ile Pro Gly Phe Gly Leu Ala Trp
145                 150                 155                 160
Thr Glu Tyr His Leu Gln Gly Arg Gln Ala His Tyr Thr Gln Gln Asn
                165                 170                 175
Trp Gly Tyr Leu Trp His Asp Phe Asp Trp Asn Trp Thr Ala Phe Asn
            180                 185                 190
Gly Ser Asp Asn Gln Leu Tyr Pro Gly Lys Trp Trp Gly Asn Thr Phe
        195                 200                 205
His Phe Pro Tyr Leu Met Gly Glu Asp Val Asp Tyr Asn Arg Phe Glu
    210                 215                 220
Val Gln Gln Glu Met Lys Ala Trp Gly Glu Trp Ile Ile Asn Ser Val
225                 230                 235                 240
Gly Phe Ser Gly Phe Arg Met Asp Ala Ile Ala His Val Asp Thr Asp
                245                 250                 255
Phe Thr Arg Asp Trp Ile Asn His Val Gln Trp Ala Thr Ser Glu Asp
            260                 265                 270
Val Phe Phe Val Ala Glu Ala Trp Val Ser Asp Ile Asn Gly Tyr Leu
        275                 280                 285
Asp Ala Val Asn Thr Pro His Leu Arg Ala Phe Asp Phe Asn Leu Arg
```

```
    290                 295                 300
Glu Asp Phe Val Ala Leu Ser Ser Gly Gly Lys Asp Met Arg Trp Trp
305                 310                 315                 320
Gly Gly Leu Val Asn Ser Gln His Arg Asp Arg Ala Val Thr Phe Val
                325                 330                 335
Asp Asn His Asp Thr Ser Arg Ala Gly Asn Pro Tyr Gly Met Pro Gln
            340                 345                 350
Val Ile Asn Tyr Lys Asn Gln Ala Tyr Ala Tyr Ile Leu Leu Arg Glu
        355                 360                 365
His Gly Val Pro Thr Val Phe Ala Arg Asp Tyr Asp Glu Phe Gly Met
    370                 375                 380
Ala Pro Thr Leu Asp Lys Leu Ile Glu Ala Arg Arg Tyr Phe Ala Tyr
385                 390                 395                 400
Gly Pro Gly His Glu Tyr Ser Gly Asn Thr Glu Ala Val Tyr Ala Tyr
                405                 410                 415
Val Arg Glu Gly Leu Ser Thr Val Pro Gly Thr Gly Leu Val Met Leu
            420                 425                 430
Ile Ser Gly Arg Asn Trp Gly Gly Gln Gln Ser Phe Thr Ile Asn Ser
        435                 440                 445
His Gln Pro Asn Thr Thr Phe Tyr Asp Tyr Thr Gly Asn Val Ser Gly
    450                 455                 460
Thr Val Thr Thr Asn Ala Gln Gly Tyr Gly Ser Phe Pro Val Thr Met
465                 470                 475                 480
Thr Glu Ser Thr Gly Trp Ser Val Trp Val Pro Gln Ser Ser Gly Gly
                485                 490                 495
Thr Gln Pro Gly Ser Ile Thr Leu Arg Met Thr Lys Asp Val Gly Tyr
            500                 505                 510
Gly Phe Ser Leu Phe Phe Thr Gly Ser Ser Ala Glu Leu Thr Asn Trp
        515                 520                 525
Gly Gly Gly Ile Glu Gly Thr Trp Thr Ser Gly His Val Trp Glu Val
    530                 535                 540
Thr Ile Pro Asp Pro Gly Asn Phe Glu Trp Lys Thr Arg Lys Gly Pro
545                 550                 555                 560
Thr Gly Gly Ser Gly Gln Asp Trp Glu Ser Gly Ser Asn His Asn Gln
                565                 570                 575
Thr Asn Leu His Pro Ser Phe Asn Gly Gly Phe Arg Ser Arg
            580                 585                 590
```