

(19) **DANMARK**

(10) **DK/EP 2405007 T3**



(12)

Oversættelse af europæisk patentskrift

Patent- og
Varemærkestyrelsen

-
- (51) Int.Cl.: **C 12 N 9/18 (2006.01)** **C 12 N 9/10 (2006.01)** **C 12 N 15/75 (2006.01)**
- (45) Oversættelsen bekendtgjort den: **2014-02-24**
- (80) Dato for Den Europæiske Patentmyndigheds bekendtgørelse om meddelelse af patentet: **2013-12-04**
- (86) Europæisk ansøgning nr.: **11184428.8**
- (86) Europæisk indleveringsdag: **2007-01-25**
- (87) Den europæiske ansøgnings publiceringsdag: **2012-01-11**
- (62) Stamansøgningsnr: **07705681.0**
- (84) Designerede stater: **AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IS IT LI LT LU LV MC NL PL PT RO SE SI SK TR**
- (73) Patenthaver: **DuPont Nutrition Biosciences ApS, Langebrogade 1 , Postboks 17, 1001 København K, Danmark**
- (72) Opfinder: **Kolkman, Marc, Linnaeushof 25, 2341 PA Oegstgeest, Holland**
Mikkelsen, JÅ_rn Dalgaard, Engkaer 38, 2650 Hvidovre, Danmark
Loretsen, Rikke HÅ_egh, Karetmagervej 52, 8900 Randers C, Danmark
- (74) Fuldmægtig i Danmark: **NORDIC PATENT SERVICE A/S, Højbro Plads 10, 1200 København K, Danmark**
- (54) Benævnelse: **Fremstilling af en lipid acyltransferase ud af transformerede Bacillus licheniformis-celler**
- (56) Fremdragne publikationer:
WO-A1-2005/066347
WO-A2-2004/064537
NERLAND A H: "THE NUCLEOTIDE SEQUENCE OF THE GENE ENCODING GCAT FROM AEROMONAS SALMONICIDA SSP. SALMONICIDA", JOURNAL OF FISH DISEASES, OXFORD, GB, vol. 19, no. 2, 1 January 1996 (1996-01-01), pages 145-150, XP008049669, ISSN: 0140-7775, DOI: 10.1111/J.1365-2761.1996.TB00693.X

Description**REFERENCE TO RELATED APPLICATIONS**

[0001] Reference is made to the following related applications: US 2002-0009518, US 2004-0091574, WO2004/064537, WO2004/064987, WO2005/066347, WO2005/066351. United States Application Serial Number 60/764,430 filed on 2 February 2006 and WO2006/008508.

FIELD OF THE PRESENT INVENTION

[0002] The present invention relates to the production of lipid acyltransferases. In particular, methods for the production of a lipid acyltransferase by expressing a lipid acyltransferase in a *Bacillus* host cell, preferably a *B. licheniformis* host cell. In addition, the present invention relates to the use of *Bacillus* (preferably *B. licheniformis*) to express a lipid acyltransferase and to a *Bacillus* host cell, preferably a *B. licheniformis* host cell, comprising in its genome a gene encoding a lipid acyltransferase.

BACKGROUND OF THE PRESENT INVENTION

[0003] Lipid acyltransferases are known to be advantageous in food applications. Lipid acyltransferases have been found to have significant acyltransferase activity in foodstuffs. This activity has surprising beneficial applications in methods of preparing foodstuffs.

[0004] For instance, WO 2004/064537 discloses a method for the *in situ* production of an emulsifier by use of a lipid acyltransferase and the advantages associated therewith. Accordingly, there is a need for a method for the commercial production of lipid acyltransferases.

[0005] However, generally genes can be difficult to express in heterologous hosts and expression of lipid acyltransferases in host cells can be problematic.

[0006] WO 2004/064537 discloses the expression of two *Aeromonas* lipid acyltransferases in *Bacillus subtilis* and *Escherichia Coli*. However, expression in *B. subtilis* is low whilst *E. coli* is not a GRAS organism and is, therefore, unsuitable as a host for enzymes that are to be used in the food industry.

[0007] US 6,255,076 discloses a method of producing a polypeptide in a *Bacillus* host cell. However, such a

method requires the use of a tandem promoter in which each promoter sequence is operably linked to a single copy of a nucleic acid sequence encoding the polypeptide sequence. Thus, there is a need in the art for an improved method for the production of lipid acyltransferases.

SUMMARY ASPECTS OF THE PRESENT INVENTION

5 [0008] Aspects of the present invention are presented in the claims and in the following commentary.

[0009] One aspect of the present invention relates to a method for the production of a lipid acyltransferase comprising the steps of:

1. (i) providing a *Bacillus* host cell wherein the *Bacillus* host cell is a *Bacillus licheniformis* cell;
- 10 2. (ii) transforming the *Bacillus* host cell wherein the *Bacillus* host cell is the *Bacillus licheniformis* cell, with a heterologous nucleotide sequence encoding a lipid acyltransferase shown as SEQ ID No. 49 and
3. (iii) expressing the lipid acyltransferase in the cell under the control of a promoter sequence.

15 [0010] A *Bacillus* host cell wherein the *Bacillus* host cell is one other than *Bacillus subtilis*, preferably a *Bacillus licheniformis* host cell, comprising a heterologous lipid acyltransferase is disclosed herein.

[0011] The use of a *Bacillus* host cell wherein the *Bacillus* host cell is one other than *Bacillus subtilis*, preferably a *Bacillus licheniformis* host cell, in the production of a heterologous lipid acyltransferase is also disclosed herein.

20 [0012] Suitably expression in the *Bacillus* host wherein the *Bacillus* host is one other than *Bacillus subtilis*, and preferably wherein the *Bacillus* host is *B. licheniformis*, may result in increased expression when compared to expression in *B. subtilis*.

25 [0013] An expression vector comprising a nucleotide sequence encoding a lipid acyltransferase operably linked to one or more regulatory sequence(s) such that the regulatory sequence(s) is capable of expressing the nucleotide sequence encoding a lipid acyltransferase in a suitable host or host cell, preferably in a *Bacillus* host (or cell) wherein the *Bacillus* host (or cell) is one other than *Bacillus subtilis*, preferably in *B. licheniformis* or a *B. licheniformis* cell is disclosed herein.

30 [0014] Suitably the lipid acyltransferase may be a recombinant lipid acyltransferase.

DETAILED ASPECTS OF THE PRESENT INVENTION

[0015] According to a first aspect of the present invention there is provided a method for the production of a lipid acyltransferase comprising the steps of:

1. (i) providing a *Bacillus* host cell wherein the *Bacillus* host cell is a *Bacillus licheniformis* cell;
- 5 2. (ii) transforming the *Bacillus* host cell wherein the *Bacillus* host cell is a *Bacillus licheniformis* cell, with an heterologous nucleotide sequence encoding a lipid acyltransferase shown as SEQ ID No. 49; and
3. (iii) expressing the lipid acyltransferase in the cell under the control of a promoter sequence.

[0016] Additionally, a nucleotide sequence encoding a signal peptide may be operably linked to said
10 heterologous nucleotide sequence encoding a lipid acyltransferase.

[0017] Suitably the method of the present invention may further comprise the additional step of isolating/recovering the lipid acyltransferase.

15 [0018] Suitably the lipid acyltransferase may be a recombinant lipid acyltransferase.

[0019] Suitably the promoter sequence used in accordance with the host cells, vectors, methods and/or uses of the present invention may be homologous to the host cell. "Homologous to the host cell" means originating within the host organism; i.e. a promoter sequence which is found naturally in the host organism. Suitably, the
20 promoter sequence may be selected from the group consisting of a nucleotide sequence encoding: an α -amylase promoter, a protease promoter, a subtilisin promoter, a glutamic acid-specific protease promoter and a levansucrase promoter. Suitably the promoter sequence may be a nucleotide sequence encoding: the LAT (e.g. the alpha-amylase promoter from *B. licheniformis*, also known as AmyL), AprL (e.g. subtilisin Carlsberg promoter), EndoGluC (e.g. the glutamic-acid specific promoter from *B. licheniformis*), AmyQ (e.g. the alpha
25 amylase promoter from *B. amyloliquefaciens* alpha-amylase promoter) and SacB (e.g. the *B. subtilis* levansucrase promoter).

[0020] In one embodiment of the present disclosure the promoter sequence is the -35 to -10 sequence of an alpha amylase promoter, preferably the -35 to -10 sequence of a *B. licheniformis* α -amylase promoter. The "-
30 35 to -10 sequence" describes the position relative to the transcription start site. Both the "-35" and the "-10" are boxes, i.e. a number of nucleotides, each comprising 6 nucleotides and these boxes are separated by 17 nucleotides. These 17 nucleotides are often referred to as a "spacer". This is illustrated in Figure 55, where the -35 and the -10 boxes are underlined. For the avoidance of doubt, where "-35 to -10 sequence" is used herein

it refers to a sequence from the start of the -35 box to the end of the -10 box i.e. including both the -35 box, the 17 nucleotide long spacer and the -10 box.

[0021] A nucleotide sequence encoding a lipid acyltransferase for use in any one of the host cells, vectors, methods and/or uses disclosed herein may comprise a GDSx motif and/ or a GANDY motif.

[0022] Preferably, the lipid acyltransferase enzyme is characterised as an enzyme which possesses acyltransferase activity and which comprises the amino acid sequence motif GDSX, wherein X is one or more of the following amino acid residues L, A, V, I, F, Y, H, Q, T, N, M or S.

[0023] Suitably, the nucleotide sequence encoding a lipid acyltransferase for use in any one of the vectors, host cells, methods and/or uses disclosed herein may be obtainable, preferably obtained, from an organism from one or more of the following genera: *Aeromonas*, *Streptomyces*, *Saccharomyces*, *Lactococcus*, *Mycobacterium*, *Streptococcus*, *Lactobacillus*, *Desulfitobacferium*, *Bacillus*, *Campylobacter*, *Vibrionaceae*, *Xylella*, *Sulfolobus*, *Aspergillus*, *Schizosaccharomyces*, *Listeria*, *Neisseria*, *Mesorhizobium*, *Ralstonia*, *Xanthomonas* and *Candida*. Preferably, the lipid acyltransferase is obtainable, preferably obtained, from an organism from the genus *Aeromonas*.

[0024] A nucleotide sequence encoding a lipid acyltransferase for use in any one of the vectors, host cells, methods and/or uses disclosed herein encodes a lipid acyltransferase that comprises an aspartic acid residue at a position corresponding to N-80 in the amino acid sequence of the *Aeromonas hydrophila* lipid acyltransferase shown as SEQ ID No. 35.

[0025] A nucleotide sequence encoding a lipid acyltransferase for use in any one of the vectors, host cells, methods and/or uses disclosed herein encodes a lipid acyltransferase that may comprise the amino acid sequence shown as SEQ ID No. 16, or an amino acid sequence which has 75% or more homology thereto. Suitably, the nucleotide sequence encoding a lipid acyltransferase encodes a lipid acyltransferase that may comprise the amino acid sequence shown as SEQ ID No. 16.

[0026] The term "heterologous" as used herein means a sequence derived from a separate genetic source or species. A heterologous sequence is a non-host sequence, a modified sequence, a sequence from a different host cell strain, or a homologous sequence from a different chromosomal location of the host cell.

[0027] A "homologous" sequence is a sequence that is found in the same genetic source or species i.e. it is naturally occurring in the relevant species of host cell.

[0028] The term "recombinant lipid acyltransferase" as used herein means that the lipid acyltransferase has been produced by means of genetic recombination. For instance, the nucleotide sequence encoding the lipid acyltransferase has been inserted into a cloning vector, resulting in a *B. licheniformis* cell characterised by the presence of the heterologous lipid acyltransferase.

5 HOST CELL

[0029] In one embodiment of the present invention the host cell for use in the methods and/or uses of the present invention is a *Bacillus licheniformis* host cell.

10 [0030] It has been found that the use of a *Bacillus licheniformis* host cell results in increased expression of a lipid acyltransferase when compared with other organisms, such as *Bacillus subtilis*.

15 [0031] A lipid acyltransferase from *Aeromonas salmonicida* has been inserted into a number of conventional expression vectors, designed to be optimal for the expression in *Bacillus subtilis*, *Hansenula polymorpha*, *Schizosaccharomyces pombe* and *Aspergillus tubigensis*, respectively. Only very low levels were, however, detected in *Hansenula polymorpha*, *Schizosaccharomyces pombe* and *Aspergillus tubigensis*. The expression levels were below 1 µg/ml, and it was not possible to select cells which yielded enough protein to initiate a commercial production (results not shown). In contrast, *Bacillus licheniformis* was able to produce protein levels, which are attractive for an economically feasible production.

20 [0032] In particular, it has been found that expression in *B. licheniformis* is approximately 100-times greater than expression in *B. subtilis* under the control of aprE promoter or is approximately 100-times greater than expression in *S. lividans* under the control of an A4 promoter and fused to cellulose (results not shown herein).

25 [0033] In another embodiment the host cell may be any *Bacillus licheniformis*.

30 [0034] The term "host cell" - in relation to the present invention includes any cell that comprises either a nucleotide sequence encoding a lipid acyltransferase as defined herein or an expression vector as described above and which is used in the recombinant production of a lipid acyltransferase having the specific properties as defined herein.

[0035] Thus, a further embodiment of the present disclosure provides a host cell comprising (for example transformed or transfected with) a nucleotide sequence of the present disclosure or a nucleotide sequence that expresses a polypeptide having the specific properties as defined herein.

[0036] Suitably, in some embodiments, the host cell may be a protease deficient or protease minus strain and/or an α -amylase deficient or α -amylase minus strain.

REGULATORY SEQUENCES

[0037] In some applications, a lipid acyltransferase sequence for use in any one of the host cells, vectors, methods and/or uses of the present disclosure may be operably linked to a regulatory sequence which is capable of providing for the expression of the nucleotide sequence, such as by the chosen host cell (such as a *B. licheniformis* cell).

[0038] By way of example, the present disclosure covers a vector comprising the nucleotide sequence of the present disclosure operably linked to such a regulatory sequence, i.e. the vector is an expression vector.

[0039] The term "operably linked" refers to a juxtaposition wherein the components described are in a relationship permitting them to function in their intended manner. A regulatory sequence "operably linked" to a coding sequence is ligated in such a way that expression of the coding sequence is achieved under conditions compatible with the control sequences.

[0040] The term "regulatory sequences" includes promoters and enhancers and other expression regulation signals.

[0041] The term "promoter" is used in the normal sense of the art, e.g. an RNA polymerase binding site.

[0042] Enhanced expression of the nucleotide sequence encoding the enzyme having the specific properties as defined herein may also be achieved by the selection of regulatory regions, e.g. promoter, secretion leader and terminator regions that are not regulatory regions for the nucleotide sequence encoding the enzyme in nature.

[0043] Suitably, the nucleotide sequence of the present disclosure may be operably linked to at least a promoter.

[0044] Suitably, the nucleotide sequence encoding a lipid acyltransferase may be operably linked to a nucleotide sequence encoding a terminator sequence. Examples of suitable terminator sequences for use in any one of the vectors, host cells, methods and/or uses of the present invention include: an α -amylase terminator sequence (for instance, CGGGACTTACCGAAAGAAACCATCAATGATGGTTTCTTTTTTGTTTCATAAA - SEQ ID No. 64), an

alkaline protease terminator sequence (for instance,

CAAGACTAAAGACCGTTCGCCCCGTTTTTGCAATAAGCGGGCGAATCTTACATAAAA ATA - SEQ ID No. 65), a glutamic-acid specific terminator sequence (for instance,

ACGGCCGTTAGATGTGACAGCCCCGTTCCAAAAGGAAGCGGGCTGTCTTCGTGTAT TATTGT -

5 SEQ ID No. 66), a levanase terminator sequence (for instance,

TCTTTTAAAGGAAAGGCTGGAATGCCCGGCA17CCAGCCACATGATCATCG7TT - SEQ ID No. 67)

and a subtilisin E terminator sequence (for instance,

GCTGACAAATAAAAAGAAGCAGGTATGGAGGAACCTGCTTCTTTTACTATTATTG). Suitably, the nucleotide sequence encoding a lipid acyltransferase may be operably linked to an α -amylase terminator, such

10 as a *B. licheniformis* α -amylase terminator.

PROMOTER

[0045] The promoter sequence to be used in accordance with the present invention may be heterologous or homologous to the sequence encoding a lipid acyltransferase.

15 [0046] The promoter sequence may be any promoter sequence capable of directing expression of a lipid acyltransferase in the host cell of choice.

[0047] Suitably, the promoter sequence may be homologous to a *Bacillus* species, for example *B. licheniformis*. Preferably, the promoter sequence is homologous to the host cell of choice.

20

[0048] Suitable promoter sequences for use in the present invention include: the promoter of the *Bacillus licheniformis* alpha-amylase gene, the promoter of the *Bacillus licheniformis* subtilisin gene, the promoter of the *Bacillus subtilis* subtilisin gene, the promoter of the *Bacillus licheniformis* alkaline protease gene (subtilisin Carlsberg gene), the promoter of the *B. licheniformis* glutamic-acid specific protease gene, the promoter of *B. amyloliquefaciens* alpha-amylase gene; the promoter of *B. subtilis* levansucrase and a "consensus" promoter having the sequence TTGACA for the "-35" region and TATAAT for the "-10" region (i.e. the -35 to -10 promoter) of the alpha-amylase gene.

25

[0049] Other examples of promoters suitable for directing the transcription of a nucleic acid sequence in the methods of the present invention include: the promoter of the *Bacillus lentus* alkaline protease gene (aprH), ; the promoter of the *Bacillus subtilis* alpha-amylase gene (amyE); the promoter of the *Bacillus stearothermophilus* maltogenic amylase gene (amyM); the promoter of the *Bacillus licheniformis* penicillinase gene (penP); the promoters of the *Bacillus subtilis* xyla and xylB genes; and/or the promoter of the *Bacillus thuringiensis* subsp. *tenebrionis* CryIIIA gene.

30

[0050] In a preferred embodiment, the promoter sequence is an α -amylase promoter (such as a *Bacillus licheniformis* α -amylase promoter). Preferably, the promoter sequence comprises the -35 to -10 sequence of the *B. licheniformis* α -amylase promoter - see Figures 53 and 55.

SIGNAL PEPTIDE

5 [0051] The lipid acyltransferase produced by a host cell by expression of the nucleotide sequence encoding the lipid acyltransferase may be secreted or may be contained intracellularly depending on the sequence and/or the vector used.

10 [0052] A signal sequence may be used to direct secretion of the coding sequences through a particular cell membrane. The signal sequences may be natural or foreign to the lipid acyltransferase coding sequence. For instance, the signal peptide coding sequence may be obtained from an amylase or protease gene from a *Bacillus* species, preferably from *Bacillus licheniformis*.

15 [0053] Suitable signal peptide coding sequences may be obtained from one or more of the following genes: maltogenic α -amylase gene, subtilisin gene, beta-lactamase gene, neutral protease gene, prsA gene, and/or acyltransferase gene.

20 [0054] Preferably, the signal peptide is a signal peptide of *B. licheniformis* α -amylase, *Aeromonas* acyltransferase (for instance, mkkwfvclgialtvqa - SEQ ID No. 21), *B. subtilis* subtilisin (for instance, mrskkiwisilfaltlftmafsnmsaqa - SEQ ID No. 22) or *B. licheniformis* subtilisin (for instance, mmrkkstwfgmaltafmlvftmefdsasa - SEQ ID No. 23). Suitably, the signal peptide may be the signal peptide of *B. licheniformis* α -amylase.

25 [0055] However, any signal peptide coding sequence capable of directing the expressed lipid acyltransferase into the secretory pathway of a *Bacillus* host cell (preferably a *B. licheniformis* host cell) of choice may be used.

[0056] In some embodiments of the present disclosure, a nucleotide sequence encoding a signal peptide may be operably linked to a nucleotide sequence encoding a lipid acyltransferase of choice.

30

[0057] The lipid acyltransferase of choice may be expressed in a host cell as defined herein as a fusion protein.

EXPRESSION VECTOR

[0058] The term "expression vector" means a construct capable of *in vivo* or *in vitro* expression.

[0059] Preferably, the expression vector is incorporated in the genome of the organism, such as a *B. licheniformis* host. The term "incorporated" preferably covers stable incorporation into the genome.

5

[0060] The nucleotide sequence encoding a lipid acyltransferase as defined herein may be present in a vector, in which the nucleotide sequence is operably linked to regulatory sequences such that the regulatory sequences are capable of providing the expression of the nucleotide sequence by a suitable host organism (such as *B. licheniformis*), i.e. the vector is an expression vector.

10

[0061] The vectors of the present invention may be transformed into a suitable host cell as described above to provide for expression of a polypeptide having lipid acyltransferase activity as defined herein.

15

[0062] The choice of vector, e.g. plasmid, cosmid, virus or phage vector, genomic insert, will often depend on the host cell into which it is to be introduced. The present invention may cover other forms of expression vectors which serve equivalent functions and which are, or become, known in the art.

20

[0063] Once transformed into the host cell of choice, the vector may replicate and function independently of the host cell's genome, or may integrate into the genome itself.

[0064] The vectors may contain one or more selectable marker genes - such as a gene which confers antibiotic resistance e.g. ampicillin, kanamycin, chloramphenicol or tetracyclin resistance. Alternatively, the selection may be accomplished by co-transformation (as described in WO91/17243).

25

[0065] Vectors may be used *in vitro*, for example for the production of RNA or used to transfect or transform a host cell.

30

[0066] Thus, in a further embodiment, the disclosure provides a method of making nucleotide sequences of the present invention or nucleotide sequences encoding polypeptides having the specific properties as defined herein for use in any one of the vectors, host cells, other methods and/or uses of the present invention, by introducing a nucleotide sequence into a replicable vector, introducing the vector into a compatible host cell, and growing the host cell under conditions which bring about replication of the vector.

[0067] The vector may further comprise a nucleotide sequence enabling the vector to replicate in the host cell

in question. Examples of such sequences are the origins of replication of plasmids pUC19, pACYC177, pUB110, pE194, AMB1 and pIJ702.

LIPID ACYL TRANSFERASE

[0068] The nucleotide sequence encoding a lipid acyl transferase for use in any one of the methods, vectors and/or uses of the present disclosure may encode a natural lipid acyl transferase or a variant lipid acyl transferase.

[0069] For instance, the nucleotide sequence encoding a lipid acyl transferase for use in the present disclosure may be one as described in WO2004/064537, WO2004/064987, WO2005/066347, or WO2006/008508.

[0070] The term "lipid acyl transferase" as used herein preferably means an enzyme that has acyltransferase activity (generally classified as E.C. 2.3.1.x, for example 2.3.1.43), whereby the enzyme is capable of transferring an acyl group from a lipid to one or more acceptor substrates, such as one or more of the following: a sterol; a stanol; a carbohydrate; a protein; a protein subunit; a sugar alcohol, such as ascorbic acid and/or glycerol - preferably glycerol and/or a sterol, such as cholesterol.

[0071] Preferably, the nucleotide sequence encoding a lipid acyl transferase for use in any one of the vectors, host cells, methods and/or uses of the present invention encodes a lipid acyltransferase that is capable of transferring an acyl group from a phospholipid (as defined herein) to a sugar alcohol, such as ascorbic acid and/or glycerol, most preferably glycerol.

[0072] For some aspects the "acyl acceptor" according to the present invention may be any compound comprising a hydroxy group (-OH), such as for example, polyvalent alcohols, including glycerol; sterols; stanols; carbohydrates; hydroxy acids including fruit acids, citric acid, tartaric acid, lactic acid and ascorbic acid; proteins or a sub-unit thereof, such as amino acids, protein hydrolysates and peptides (partly hydrolysed protein) for example; and mixtures and derivatives thereof. Preferably, the "acyl acceptor" according to the present invention is not water. Preferably, the "acyl acceptor" according to the present invention is a sugar alcohol, such as a polyol, most preferably glycerol. For the purpose of this invention ascorbic acid is also considered a sugar-alcohol.

[0073] The acyl acceptor is preferably not a monoglyceride.

[0074] The acyl acceptor is preferably not a diglyceride

[0075] In one aspect, the nucleotide sequence encoding a lipid acyltransferase for use in any one of the host cells, vectors, methods and/or uses of the present invention encodes a lipid acyltransferase that may, as well as being able to transfer an acyl group from a lipid to glycerol, additionally be able to transfer the acyl group from a lipid to one or more of the following: a carbohydrate, a protein, a protein subunit, sterol and/or a stanol, preferably it is capable of transferring to both a sugar alcohol, such as ascorbic acid and/or glycerol, most preferably a sterol such as cholesterol, and/or plant sterol/stanols.

[0076] Preferably, the lipid substrate upon which the lipid acyl acts is one or more of the following lipids: a phospholipid, such as a lecithin, e.g. phosphatidylcholine.

[0077] This lipid substrate may be referred to herein as the "lipid acyl donor". The term lecithin as used herein encompasses phosphatidylcholine, phosphatidylethanolamine, phosphatidylinositol, phosphatidylserine and phosphatidylglycerol.

[0078] For some aspects, preferably the nucleotide sequence encoding a lipid acyl transferase for use in any one of the host cells, vectors, methods and/or uses of the present invention encodes a lipid acyltransferase that is incapable, or substantially incapable, of acting on a triglyceride and/or a 1-monoglyceride and/or 2-monoglyceride.

[0079] For some aspects, preferably the nucleotide sequence encoding a lipid acyl transferase for use in any one of the host cells, vectors, methods and/or uses of the present invention encodes a lipid acyltransferase that does not exhibit triacylglycerol lipase activity (E.C. 3.1.1.3) or does not exhibit significant triacylglycerol lipase activity (E.C. 3.1.1.3).

[0080] The ability to hydrolyse triglyceride (E.C. 3.1.1.3 activity) may be determined by lipase activity is determined according to Food Chemical Codex (3rd Ed., 1981, pp 492-493) modified to sunflower oil and pH 5.5 instead of olive oil and pH 6.5. The lipase activity is measured as LUS (lipase units sunflower) where 1 LUS is defined as the quantity of enzyme which can release 1 [mu]mol of fatty acids per minute from sunflower oil under the above assay conditions. Alternatively the LUT assay as defined in WO9845453 may be used.

[0081] The nucleotide sequence encoding a lipid acyl transferase for use in any one of the host cells, vectors, methods and/or uses of the present invention may encode a lipid acyltransferase that which is substantially incapable of acting on a triglyceride may have a LUS/mg of less than 1000, for example less than 500, such as

less than 300, preferably less than 200, more preferably less than 100, more preferably less than 50, more preferably less than 20, more preferably less than 10, such as less than 5, less than 2, more preferably less than 1 LUS/mg. Alternatively LUT/mg activity is less than 500, such as less than 300, preferably less than 200, more preferably less than 100, more preferably less than 50, more preferably less than 20, more preferably less than 10, such as less than 5, less than 2, more preferably less than 1 LUT/mg.

[0082] The nucleotide sequence encoding a lipid acyl transferase for use in any one of the host cells, vectors, methods and/or uses of the present invention may encode a lipid acyltransferase that which is substantially incapable of acting on a monoglyceride may be determined by using mono-oleate (M7765 1-Oleoyl-rac-glycerol 99%) in place of the sunflower oil in the LUS assay. 1 MGHU is defined as the quantity of enzyme which can release 1 [mu]mol of fatty acids per minute from monoglyceride under the assay conditions.

[0083] The nucleotide sequence encoding a lipid acyl transferase for use in any one of the host cells, vectors, methods and/or uses of the present invention encodes a lipid acyltransferase that which is substantially incapable of acting on a triglyceride may have a MGHU/mg of less than 5000, for example less than 1000, for example less than 500, such as less than 300, preferably less than 200, more preferably less than 100, more preferably less than 50, more preferably less than 20, more preferably less than 10, such as less than 5, less than 2, more preferably less than 1 MGHU/mg.

[0084] Suitably, the nucleotide sequence encoding a lipid acyltransferase for use in any one of the host calls, vectors, methods and/or uses of the present invention encodes a lipid acyltransferase that may exhibit one or more of the following phospholipase activities: phospholipase A2 activity (E.C. 3.1.1.4) and/or phospholipase A1 activity (E.C. 3.1.1.32). The lipid acyl transferase may also have phospholipase B activity (E.C 3.1.1.5).

[0085] Suitably, for some aspects the lipid acyltransferase may be capable of transferring an acyl group from a phospholipid to a sugar alcohol, preferably glycerol and/or ascorbic acid.

[0086] For some aspects, preferably the nucleotide sequence encoding a lipid acyltransferase for use any one of the host cells, vectors, methods and/or uses of the present invention encodes a lipid acyltransferase that is capable of transferring an acyl group from a phospholipid to a sterol and/or a stanol to form at least a sterol ester and/or a stanol ester.

[0087] The lipid acyltransferase may be capable of transferring an acyl group from a lipid to a polyol such as glycerol, and/or a sterol such as cholesterol or plant sterol/stanols. Thus, in one embodiment the "acyl acceptor" according to the present invention may be glycerol and/or cholesterol or plant sterol/stanols.

[0088] Preferably, the lipid acyltransferase enzyme may be characterised using the following criteria:

the enzyme possesses acyl transferase activity which may be defined as ester transfer activity whereby the acyl part of an original ester bond of a lipid acyl donor is transferred to an acyl acceptor, preferably glycerol or cholesterol, to form a new ester; and

- 5 the enzyme comprises the amino acid sequence motif GDSX, wherein X is one or more of the following amino acid residues L, A, V, I, F, Y, H, Q, T, N, M or S.

[0089] Preferably, X of the GDSX motif is L or Y. More preferably, X of the GDSX motif is L. Thus, preferably the enzyme according to the present invention comprises the amino acid sequence motif GDSL.

- 10 [0090] The GDSX motif is comprised of four conserved amino acids. Preferably, the serine within the motif is a catalytic serine of the lipid acyl transferase enzyme. Suitably, the serine of the GDSX motif may be in a position corresponding to Ser-16 in *Aeromonas. hydrophila* lipid acyltransferase enzyme taught in Brumlik & Buckley (Journal of Bacteriology Apr. 1996, Vol. 178, No. 7, p 2060-2064).

- 15 [0091] To determine if a protein has the GDSX motif according to the present invention, the sequence is preferably compared with the hidden markov model profiles (HMM profiles) of the pfam database in accordance with the procedures taught in WO2004/064537 or WO20041064987.

- 20 [0092] Preferably the lipid acyl transferase enzyme can be aligned using the Pfam00657 consensus sequence (for a full explanation see WO2004/064537 or WO2004/064987).

[0093] Preferably, a positive match with the hidden markov model profile (HMM profile) of the pfam00657 domain family indicates the presence of the GDSL or GDSX domain according to the present invention.

- 25 [0094] Preferably when aligned with the Pfam00657 consensus sequence the lipid acyltransferase for use in the methods or uses of the invention may have at least one, preferably more than one, preferably more than two, of the following, a GDSx block, a GANDY block, a HPT block. Suitably, the lipid acyltransferase may have a GDSx block and a GANDY block. Alternatively, the enzyme may have a GDSx block and a HPT block. Preferably the enzyme comprises at least a GDSx block. See WO2004/064537 or WO2004/064987 for
30 further details.

[0095] Preferably, residues of the GANDY motif are selected from GANDY, GGND A, GGNDL, most preferably GANDY.

[0096] Preferably, when aligned with the Pfam00657 consensus sequence the enzyme for use in the methods or uses of the invention have at least one, preferably more than one, preferably more than two, preferably more than three, preferably more than four, preferably more than five, preferably more than six, preferably more than seven, preferably more than eight, preferably more than nine, preferably more than ten, preferably more than eleven, preferably more than twelve, preferably more than thirteen, preferably more than fourteen, of the following amino acid residues when compared to the reference *A. hydrophila* polypeptide sequence, namely SEQ ID No. 1: 28hid, 29hid, 30hid, 31hid, 32gly, 33Asp, 34Ser, 35hid, 130hid, 131Gly, 132Hid, 133Asn, 134Asp, 135hid, 309His.

[0097] The pfam00657 GDSX domain is a unique identifier which distinguishes proteins possessing this domain from other enzymes.

[0098] The pfam00657 consensus sequence is presented in Figure 3 as SEQ ID No. 2. This is derived from the identification of the pfam family 00657, database version 6, which may also be referred to as pfam00657.6 herein.

[0099] The consensus sequence may be updated by using further releases of the pfam database (for example see WO2004/064537 or WO2004/064987).

[0100] In one embodiment, the nucleotide sequence encoding a lipid acyl transferase enzyme for use in any one of the host cells, vectors, methods and/or uses of the present invention encodes a lipid acyltransferase that may be characterised using the following criteria:

1. (i) the enzyme possesses acyl transferase activity which may be defined as ester transfer activity whereby the acyl part of an original ester bond of a lipid acyl donor is transferred to acyl acceptor, preferably glycerol or cholesterol, to form a new ester, preferably monoglyceride or cholesterol ester respectfully;
2. (ii) the enzyme comprises the amino acid sequence motif GDSX, wherein X is one or more of the following amino acid residues L, A, V, I, F, Y, H, Q, T, N, M or S.;
3. (iii) the enzyme comprises His-309 or comprises a histidine residue at a position corresponding to His-309 in the *Aeromonas hydrophila* lipid acyltransferase enzyme shown in Figures 2 and 4 (SEQ ID No. 1 or SEQ ID No. 3).

[0101] Preferably, the amino acid residue of the GDSX motif is L.

[0102] In SEQ ID No. 3 or SEQ ID No. 1 the first 18 amino acid residues form a signal sequence. His-309 of the full length sequence, that is the protein including the signal sequence, equates to His-291 of the mature part of the protein, i.e. the sequence without the signal sequence.

5 [0103] The nucleotide sequence encoding a lipid acyl transferase enzyme for use any one of the host cells, vectors, methods and uses disclosed herein encodes a lipid acyltransferase that comprises the following catalytic triad: Ser-34, Asp-306 and His-309 or comprises a serine residue, an aspartic acid residue and a histidine residue, respectively, at positions corresponding to Ser-34, Asp-306 and His-309 in the *Aeromonas hydrophila* lipid acyl transferase enzyme shown in Figure 4 (SEQ ID No. 3) or Figure 2 (SEQ ID No. 1). As
10 stated above, in the sequence shown in SEQ ID No. 3 or SEQ ID No. 1 the first 18 amino acid residues form a signal sequence. Ser-34, Asp-306 and His-309 of the full length sequence, that is the protein including the signal sequence, equate to Ser-16, Asp-288 and His-291 of the mature part of the protein, i.e. the sequence without the signal sequence. In the pfam00657 consensus sequence, as given in Figure 3 (SEQ ID No. 2) the active site residues correspond to Ser-7, Asp-345 and His-348.

15 [0104] The nucleotide sequence encoding a lipid acyl transferase enzyme for use any one of the host cells, vectors, methods and uses disclosed herein encodes a lipid acyltransferase that may be characterised using the following criteria:

20 the enzyme possesses acyl transferase activity which may be defined as ester transfer activity whereby the acyl part of an original ester bond of a first lipid acyl donor is transferred to an acyl acceptor to form a new ester; and

25 the enzyme comprises at least Gly-32, Asp-33, Ser-34, Asp-134 and His-309 or comprises glycine, aspartic acid, serine, aspartic acid and histidine residues at positions corresponding to Gly-32, Asp-33, Ser-34, Asp-306 and His-309, respectively, in the *Aeromonas hydrophila* lipid acyltransferase enzyme shown in SEQ ID No. 3 or SEQ ID No. 1.

[0105] The nucleotide sequence encoding a lipid acyltransferase enzyme for use in any one of the host cells, vectors, methods and uses disclosed herein may be one of the following nucleotide sequences:

1. (a) the nucleotide sequence shown as SEQ ID No. 36 (see Figure 29);
2. (b) the nucleotide sequence shown as SEQ ID No. 38 (see Figure 31);
- 30 3. (c) the nucleotide sequence shown as SEQ ID No. 39 (see Figure 32);
4. (d) the nucleotide sequence shown as SEQ ID No. 42 (see Figure 35);
5. (e) the nucleotide sequence shown as SEQ ID No. 44 (see Figure 37);

6. (f) the nucleotide sequence shown as SEQ ID No. 46 (see Figure 39);
7. (g) the nucleotide sequence shown as SEQ ID No. 48 (see Figure 41);
8. (h) the nucleotide sequence shown as SEQ ID No. 49 (see Figure 57);
9. (i) the nucleotide sequence shown as SEQ ID No. 50 (see Figure 58);
- 5 10. (j) the nucleotide sequence shown as SEQ ID No. 51 (see Figure 59);
11. (k) the nucleotide sequence shown as SEQ ID No. 52 (see Figure 60);
12. (l) the nucleotide sequence shown as SEQ ID No. 53 (see Figure 61);
13. (m) the nucleotide sequence shown as SEQ ID No. 54 (see Figure 62);
14. (n) the nucleotide sequence shown as SEQ ID No. 55 (see Figure 63);
- 10 15. (o) the nucleotide sequence shown as SEQ ID No. 56 (see Figure 64);
16. (p) the nucleotide sequence shown as SEQ ID No. 57 (see Figure 65);
17. (q) the nucleotide sequence shown as SEQ ID No. 58 (see Figure 66);
18. (r) the nucleotide sequence shown as SEQ ID No. 59 (see Figure 67);
19. (s) the nucleotide sequence shown as SEQ ID No. 60 (see Figure 68);
- 15 20. (t) the nucleotide sequence shown as SEQ ID No. 61 (see Figure 69);
21. (u) the nucleotide sequence shown as SEQ ID No. 62 (see Figure 70);
22. (v) the nucleotide sequence shown as SEQ ID No. 63 (see Figure 71);
23. (w) or

a nucleotide sequence which has 70% or more, preferably 75% or more, identity with any one of the

20 sequences shown as SEQ ID No. 36, SEQ ID No. 38, SEQ ID No. 39, SEQ ID No. 42, SEQ ID No. 44, SEQ ID No. 46, SEQ ID No. 48, SEQ ID No. 49, SEQ ID No. 50, SEQ ID No. 51, SEQ ID No. 52, SEQ ID No. 53, SEQ ID No. 54, SEQ ID No. 55, SEQ ID No. 56, SEQ ID No. 57, SEQ ID No. 58, SEQ ID No. 59, SEQ ID No. 60, SEQ ID No. 61, SEQ ID No. 62 or SEQ ID No. 63.

25 **[0106]** Suitably the nucleotide sequence may have 80% or more, preferably 85% or more, more preferably 90% or more and even more preferably 95% or more identity with any one of the sequences shown as SEQ ID No. 36, SEQ ID No. 38, SEQ ID No. 39, SEQ ID No. 42, SEQ ID No. 44, SEQ ID No. 46, SEQ ID No. 48, SEQ ID No. 49, SEQ ID No. 50, SEQ ID No. 51, SEQ ID No. 52, SEQ ID No. 53, SEQ ID No. 54, SEQ ID No. 55, SEQ ID No. 56, SEQ ID No. 57, SEQ ID No. 58, SEQ ID No. 59, SEQ ID No. 60, SEQ ID No. 61, 30 SEQ ID No. 62 or SEQ ID No. 63.

[0107] The nucleotide sequence encoding a lipid acyltransferase enzyme for use any one of the host cells, vectors, methods and uses disclosed herein is a nucleotide sequence which has 70% or more, preferably 75% or more, identity with any one of the sequences shown as: SEQ ID No. 49, SEQ ID No. 50, SEQ ID No. 51,

SEQ ID No. 62, and SEQ ID No. 63. Suitably the nucleotide sequence may have 80% or more, preferably 85% or more, more preferably 90% or more and even more preferably 95% or more identity with any one of the sequences shown as: SEQ ID No. 49, SEQ ID No. 50, SEQ ID No. 51, SEQ ID No. 62, and SEQ ID No. 63.

5

[0108] In one embodiment, the nucleotide sequence encoding a lipid acyltransferase enzyme for use in any one of the host cells, vectors, methods and uses of the present disclosure is a nucleotide sequence which has 70% or more, 75% or more, 80% or more, preferably 85% or more, more preferably 90% or more and even more preferably 95% or more identity the sequence shown as SEQ ID No. 49.

10

[0109] The nucleotide sequence encoding a lipid acyl transferase enzyme for use any one of the host cells, vectors, methods and uses disclosed herein may encode a lipid acyltransferase that comprises one or more of the following amino acid sequences:

1. (i) the amino acid sequence shown as SEQ ID No. 3
- 15 2. (ii) the amino acid sequence shown as SEQ ID No. 4
3. (iii) the amino acid sequence shown as SEQ ID No. 5
4. (iv) the amino acid sequence shown as SEQ ID No. 6
5. (v) the amino acid sequence shown as SEQ ID No. 7
6. (vi) the amino acid sequence shown as SEQ ID No. 8
- 20 7. (vii) the amino acid sequence shown as SEQ ID No. 9
8. (viii) the amino acid sequence shown as SEQ ID No. 10
9. (ix) the amino acid sequence shown as SEQ ID No. 11
10. (x) the amino acid sequence shown as SEQ ID No. 12
11. (xi) the amino acid sequence shown as SEQ ID No. 13
- 25 12. (xii) the amino acid sequence shown as SEQ ID No. 14
13. (xiii) the amino acid sequence shown as SEQ ID No. 1
14. (xiv) the amino acid sequence shown as SEQ ID No. 15 or

an amino acid sequence which has 75%, 80%, 85%, 90%, 95%, 98% or more identity with any one of the sequences shown as SEQ ID No. 1, SEQ ID No. 3, SEQ ID No. 4, SEQ ID No. 5, SEQ ID No. 6, SEQ ID No. 7, SEQ ID No. 8, SEQ ID No. 9, SEQ ID No. 10, SEQ ID No. 11, SEQ ID No. 12, SEQ ID No. 13, SEQ ID No. 14, or SEQ ID No. 15.

30

[0110] A nucleotide sequence encoding a lipid acyl transferase enzyme for use any one of the host cells, vectors, methods and uses disclosed herein may encode a lipid acyltransferase that comprises either the amino

acid sequence shown as SEQ ID No. 3 or as SEQ ID No. 4 or SEQ ID No. 1 or SEQ ID No. 15 or comprises an amino acid sequence which has 75% or more, preferably 80% or more, preferably 85% or more, preferably 90% or more, preferably 95% or more, identity with the amino acid sequence shown as SEQ ID No. 3 or the amino acid sequence shown as SEQ ID No. 4 or the amino acid sequence shown as SEQ ID No. 1 or the amino acid sequence shown as SEQ ID No. 15.

[0111] The nucleotide sequence encoding a lipid acyl transferase enzyme for use any one of the host cells, vectors, methods and uses disclosed herein may encode a lipid acyltransferase that comprises an amino acid sequence which has 80% or more, preferably 85% or more, more preferably 90% or more and even more preferably 95% or more identity with any one of the sequences shown as SEQ ID No. 3, SEQ ID No. 4, SEQ ID No. 5, SEQ ID No. 6, SEQ ID No. 7, SEQ ID No. 8, SEQ ID No. 9, SEQ ID No. 10, SEQ ID No. 11, SEQ ID No. 12, SEQ ID No. 13, SEQ ID No. 14, SEQ ID No. 1, or SEQ ID No. 15.

[0112] The nucleotide sequence encoding a lipid acyl transferase enzyme for use any one of the host cells, vectors, methods and uses disclosed herein may encode a lipid acyltransferase that comprises one or more of the following amino acid sequences:

1. (a) an amino acid sequence shown as amino acid residues 1-100 of SEQ ID No. 3 or SEQ ID No. 1;
2. (b) an amino acid sequence shown as amino acids residues 101-200 of SEQ ID No. 3 or SEQ ID No. 1;
3. (c) an amino acid sequence shown as amino acid residues 201-300 of SEQ ID No. 3 or SEQ ID No. 1; or
4. (d) an amino acid sequence which has 75% or more, preferably 85% or more, more preferably 90% or more, even more preferably 95% or more identity to any one of the amino acid sequences defined in (a)-(c) above.

[0113] The lipid acyl transferase enzyme for use in methods and uses disclosed herein may comprise one or more of the following amino acid sequences:

1. (a) an amino acid sequence shown as amino acid residues 28-39 of SEQ ID No. 3 or SEQ ID No. 1;
2. (b) an amino acid sequence shown as amino acids residues 77-88 of SEQ ID No. 3 or SEQ ID No. 1;
3. (c) an amino acid sequence shown as amino acid residues 126-136 of SEQ ID No. 3 or SEQ ID No. 1;
4. (d) an amino acid sequence shown as amino acid residues 163-175 of SEQ ID No. 3 or SEQ ID No. 1;
5. (e) an amino acid sequence shown as amino acid residues 304-311 of SEQ ID No. 3 or SEQ ID No. 1; or

6. (f) an amino acid sequence which has 75% or more, preferably 85% or more, more preferably 90% or more, even more preferably 95% or more identity to any one of the amino acid sequences defined in (a)-(e) above.

[0114] The nucleotide sequence encoding a lipid acyl transferase enzyme for use any one of the host cells, vectors, methods and uses disclosed herein encodes a lipid acyltransferase that may be the lipid acyl transferase from *Candida parapsilosis* as taught in EP 1 275 711. Thus the lipid acyl transferase for use in the method and uses disclosed herein may be a lipid acyl transferase comprising one of the amino acid sequences taught in SEQ ID No. 17 or SEQ ID No. 18.

[0115] Much by preference, the nucleotide sequence encoding a lipid acyl transferase enzyme for use in any one of the host cells, vectors, methods and uses of the present invention encodes a lipid acyltransferase that may be a lipid acyl transferase (lipid acyltransferase) comprising the amino acid sequence shown as SEQ ID No. 16, or an amino acid sequence which has 75% or more, preferably 85% or more, more preferably 90% or more, even more preferably 95% or more, even more preferably 98% or more, or even more preferably 99% or more identity to SEQ ID No. 16. This enzyme could be considered a variant enzyme.

[0116] In one aspect, the nucleotide sequence encoding a lipid acyltransferase enzyme for use any one of the host cells, vectors, methods and uses of the present invention encodes a lipid acyltransferase that may be a lecithin:cholesterol acyltransferase (LCAT) or variant thereof (for example a variant made by molecular evolution)

[0117] Suitable LCATs are known in the art and may be obtainable from one or more of the following organisms for example: mammals, rat, mice, chickens, *Drosophila melanogaster*, plants, including *Arabidopsis* and *Oryza sativa*, nematodes, fungi and yeast.

[0118] In one embodiment the nucleotide sequence encoding a lipid acyltransferase enzyme for use any one of the host cells, vectors, methods and uses of the present disclosure encodes a lipid acyltransferase that may be the lipid acyltransferase obtainable, preferably obtained, from the *E. coli* strains TOP 10 harbouring pPet12aAhydro and pPet12aASalmo deposited by Danisco A/S of Langebrogade 1, DK-1001 Copenhagen K, Denmark under the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the purposes of Patent Procedure at the National Collection of Industrial, Marine and Food Bacteria (NCIMB) 23 St. Machar Street, Aberdeen Scotland, GB on 22 December 2003 under accession numbers NCIMB 41204 and NCIMB 41205, respectively.

[0119] A nucleotide sequence encoding a lipid acyltransferase enzyme for use any one of the host cells, vectors, methods and uses disclosed herein may encode a phospholipid glycerol acyl transferase. Phospholipid glycerol acyl transferases include those isolated from *Aeromonas spp.*, preferably *Aeromonas hydrophila* or *A. salmonicida*, most preferable *A. salmonicida* or variants thereof. Lipid acyl transferases disclosed herein are encoded by SEQ ID No.s 1, 3, 4, 15 and 16. It will be recognised by the skilled person that it is preferable that the signal peptides of the acyl transferase has been cleaved during expression of the transferase. The signal peptide of SEQ ID 1, 3, 4, 15 and 16 are amino acids 1-18. Therefore the most preferred regions are amino acids 19-335 for SEQ ID No. 1 and SEQ ID No. 3 (*A. hydrophila*) and amino acids 19-336 for SEQ ID No. 4, SEQ ID No. 15 and SEQ ID No. 16. (*A. salmonicida*). When used to determine the homology of identity of the amino acid sequences, it is preferred that the alignments as herein described use the mature sequence.

[0120] Therefore the most preferred regions for determining homology (identity) are amino acids 19-335 for SEQ ID No. 1 and 3 (*A. hydrophila*) and amino acids 19-336 for SEQ ID No.s 4, 15 and 16. (*A. salmonicida*). SEQ ID 34 and 35 are mature protein sequences of a lipid acyl transferase from *A. hydrophila* and *A. salmonicida* respectively.

[0121] A nucleotide sequence encoding a lipid acyltransferase enzyme for use any one of the host cells, vectors, methods and uses disclosed herein encodes a lipid acyltransferase that may also be isolated from *Thermobifida*, preferably *T. fusca*, most preferably that encoded by SEQ ID No. 28.

[0122] A nucleotide sequence encoding a lipid acyltransferase enzyme for use any one of the host cells, vectors, methods and uses disclosed herein encodes a lipid acyltransferase that may also be isolated from *Streptomyces*, preferable *S. avermitis*, most preferably that encoded by SEQ ID No. 32. Other possible enzymes from *Streptomyces* disclosed herein include those encoded by SEQ ID No.s 5, 6, 9, 10, 11, 12, 13, 14, 31, and 33.

[0123] An enzyme for uses disclosed herein may also be isolated from *Corynebacterium*, preferably *C. efficiens*, most preferably that encoded by SEQ ID No. 29.

[0124] The nucleotide sequence encoding a lipid acyltransferase enzyme for use any one of the host cells, vectors, methods and uses disclosed herein may encode a lipid acyltransferase that comprises any one of the amino acid sequences shown as SEQ ID No.s 37, 38, 40, 41, 43, 45, or 47 or an amino acid sequence which has at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97% or 98% identity therewith, or encoded by any one of the nucleotide sequences shown as SEQ ID No.s 36, 39, 42, 44, 46, or 48 or a nucleotide sequence which has at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97% or 98% identity therewith.

[0125] The nucleic sequence encoding a lipid acyltransferase enzyme for use any one of the host cells, vectors, methods and uses disclosed herein is selected from the group consisting of:

1. a) a nucleic acid comprising a nucleotide sequence shown in SEQ ID No. 36;
2. b) a nucleic acid which is related to the nucleotide sequence of SEQ ID No. 36 by the degeneration of the genetic code; and
3. c) a nucleic acid comprising a nucleotide sequence which has at least 70% identity with the nucleotide sequence shown in SEQ ID No. 36.

[0126] In one embodiment, a nucleotide sequence encoding a lipid acyltransferase enzyme for use any one of the host cells, vectors, methods and uses disclosed herein encodes a lipid acyltransferase that comprises an amino acid sequence as shown in SEQ ID No. 37 or an amino acid sequence which has at least 60% identity thereto.

[0127] The nucleotide sequence encoding a lipid acyltransferase enzyme for use any one of the host cells, vectors, methods and uses disclosed herein may encode a lipid acyltransferase comprising any one of the amino acid sequences shown as SEQ ID No. 37, 38, 40, 41, 43, 45 or 47 or an amino acid sequence which has at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97% or 98% identity therewith, or encoded by any one of the nucleotide sequences shown as SEQ ID No. 39, 42, 44, 46 or 48 or a nucleotide sequence which has at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97% or 98% identity therewith.

[0128] The nucleotide sequence encoding a lipid acyltransferase enzyme for use any one of the host cells, vectors, methods and uses disclosed herein may encode a lipid acyltransferase comprising any one of amino sequences shown as SEQ ID No. 38, 40, 41, 45 or 47 or an amino acid sequence which has at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97% or 98% identity therewith for the uses described herein.

[0129] The nucleotide sequence encoding a lipid acyltransferase for use in any one of the host cells, vectors, methods and uses disclosed herein may encode a lipid acyltransferase comprising any one of amino sequences shown as SEQ ID No. 38, 40, or 47 or an amino acid sequence which has at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97% or 98% identity therewith for the uses described herein.

[0130] The nucleotide sequence encoding a lipid acyltransferase for use in any one of the host cells, vectors, methods and uses disclosed herein may encode a lipid acyltransferase comprising the amino acid sequence shown as SEQ ID No. 47 or an amino acid sequence which has at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97% or 98% identity therewith.

[0131] The nucleotide sequence encoding a lipid acyltransferase for use in any one of the host cells, vectors, methods and uses disclosed herein may encode a lipid acyltransferase comprising the amino acid sequence shown as SEQ ID No. 43 or 44 or an amino acid sequence which has at least 80%, 85%, 90%, 95%, 96%, 97% or 98% identity therewith.

5

[0132] The nucleotide sequence encoding a lipid acyltransferase for use in any one of the host cells, vectors, methods and uses disclosed herein may encode a lipid acyltransferase comprising the amino acid sequence shown as SEQ ID No. 41 or an amino acid sequence which has at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97% or 98% identity therewith.

10

[0133] The nucleotide sequence encoding a lipid acyltransferase for use in any one of the host cells, vectors, methods and uses disclosed herein is selected from the group consisting of:

1. a) a nucleic acid comprising a nucleotide sequence shown in SEQ ID No. 36;
2. b) a nucleic acid which is related to the nucleotide sequence of SEQ ID No. 36 by the degeneration of the genetic code; and
3. c) a nucleic acid comprising a nucleotide sequence which has at least 70% identity with the nucleotide sequence shown in SEQ ID No. 36.

15

[0134] The lipid acyltransferase disclosed herein may be a lipid acyltransferase obtainable, preferably obtained, from the *Streptomyces* strains L130 or L131 deposited by Danisco A/S of Langebrogade 1, DK-1001 Copenhagen K, Denmark under the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the purposes of Patent Procedure at the National Collection of Industrial, Marine and Food Bacteria (NCIMB) 23 St. Machar Street, Aberdeen Scotland, GB on 25 June 2004 under accession numbers NCIMB 41226 and NCIMB 41227, respectively.

20

[0135] Nucleotide sequences encoding a lipid acyltransferase for use in any one of the host cells, vectors, methods and uses disclosed herein may encode a polynucleotide encoding a lipid acyltransferase (SEQ ID No. 16); or may encode an amino acid sequence of a lipid acyltransferase (SEQ ID No. 17).

25

[0136] A nucleotide sequence encoding a lipid acyltransferase for use in any one of the host cells, vectors, methods and uses disclosed herein may encode an amino acid sequence which may be identified by alignment to the L131 (SEQ ID No. 37) sequence using Align X, the Clustal W pairwise alignment algorithm of VectorNTI using default settings.

30

[0137] An alignment of the L131 and homologues from *S. avermitilis* and *T. fusca* illustrates that the

conservation of the GDS_x motif (GDSY in L131 and *S. avermitilis* and *T. fusca*), the GANDY box, which is either GGND_A or GGND_L, and the HPT block (considered to be the conserved catalytic histidine). These three conserved blocks are highlighted in Figure 42.

5 **[0138]** When aligned to either the pfam Pfam00657 consensus sequence (as described in WO04/064987) and/or the L131 sequence herein disclosed (SEQ ID No 37) it is possible to identify three conserved regions, the GDS_x block, the GANDY block and the HTP block (see WO04/064987 for further details).

10 **[0139]** When aligned to either the pfam Pfam00657 consensus sequence (as described in WO04/064987) and/or the L131 sequence herein disclosed (SEQ ID No 37)

1. i) The nucleotide sequence encoding a lipid acyltransferase for use in any one of the host cells, vectors, methods and uses disclosed herein may encode a lipid acyltransferase that, has a GDS_x motif, more preferably a GDS_x motif selected from GDSL or GDSY motif.
and/or
- 15 2. ii) The nucleotide sequence encoding a lipid acyltransferase for use in any one of the host cells, vectors, methods and uses disclosed herein may encode a lipid acyltransferase that, has a GANDY block, more preferably a GANDY block comprising amino GGND_x, more preferably GGND_A or GGND_L.
and/or
- 20 3. iii) The nucleotide sequence encoding a lipid acyltransferase for use in any one of the host cells, vectors, methods and uses disclosed herein encodes a lipid acyltransferase that has preferably an HTP block.
and preferably
- 25 4. iv) nucleotide sequence encoding a lipid acyltransferase for use in any one of the host cells, vectors, methods and uses disclosed herein may encode a lipid acyltransferase that has preferably a GDS_x or GDSY motif, and a GANDY block comprising amino GGND_x, preferably GGND_A or GGND_L, and a HTP block (conserved histidine).

Variant lipid acyl transferase

30 **[0140]** In a preferred embodiment the nucleotide sequence encoding a lipid acyltransferase for use in any one of the host cells, vectors, methods and uses of the present disclosure may encode a lipid acyltransferase that is a variant lipid acyl transferase.

[0141] Variants which have an increased activity on phospholipids, such as increased hydrolytic activity and/or increased transferase activity, preferably increased transferase activity on phospholipids may be used.

[0142] Preferably the variant lipid acyltransferase is prepared by one or more amino acid modifications of the lipid acyl transferases as defined hereinabove.

[0143] Suitably, when the nucleotide sequence encoding a lipid acyltransferase for use in any one of the host cells, vectors, methods and uses of the present disclosure may encode a lipid acyltransferase that may be a variant lipid acyltransferase, in which case the enzyme may be characterised in that the enzyme comprises the amino acid sequence motif GDSX, wherein X is one or more of the following amino acid residues L, A, V, I, F, Y, H, Q, T, N, M or S, and wherein the variant enzyme comprises one or more amino acid modifications compared with a parent sequence at any one or more of the amino acid residues defined in set 2 or set 4 or set 6 or set 7 (as defined WO2005/066347 and hereinbelow).

[0144] For instance the variant lipid acyltransferase may be characterised in that the enzyme comprises the amino acid sequence motif GDSX, wherein X is one or more of the following amino acid residues L, A, V, I, F, Y, H, Q, T, N, M or S, and wherein the variant enzyme comprises one or more amino acid modifications compared with a parent sequence at any one or more of the amino acid residues detailed in set 2 or set 4 or set 6 or set 7 (as defined in WO2005/066347 and hereinbelow) identified by said parent sequence being structurally aligned with the structural model of P10480 defined herein, which is preferably obtained by structural alignment of P10480 crystal structure coordinates with 1IVN.PDB and/or 1DEO.PDB as defined WO2005/066347 and hereinbelow.

[0145] A nucleotide sequence encoding a lipid acyltransferase for use in any one of the host cells, vectors, methods and uses disclosed herein may encode a variant lipid acyltransferase that may be characterised in that the enzyme comprises the amino acid sequence motif GDSX, wherein X is one or more of the following amino acid residues L, A, V, I, F, Y, H, Q, T, N, M or S, and wherein the variant enzyme comprises one or more amino acid modifications compared with a parent sequence at any one or more of the amino acid residues taught in set 2 identified when said parent sequence is aligned to the pfam consensus sequence (SEQ ID No. 2 -Figure 3) and modified according to a structural model of P10480 to ensure best fit overlap as defined WO2005/066347 and hereinbelow.

[0146] The nucleotide sequence encoding a lipid acyltransferase for use in any one of the host cells, vectors, methods and uses disclosed herein may encode a variant lipid acyltransferase enzyme that may comprise an amino acid sequence, which amino acid sequence is shown as SEQ ID No. 34, SEQ ID No. 3, SEQ ID No. 4,

SEQ ID No. 5, SEQ ID No. 6, SEQ ID No. 7, SEQ ID No. 8, SEQ ID No. 19, SEQ ID No. 10, SEQ ID No. 11, SEQ ID No. 12, SEQ ID No. 13, SEQ ID No. 14, SEQ ID No. 1, SEQ ID No. 15, SEQ ID No. 25, SEQ ID No. 26, SEQ ID No. 27, SEQ ID No. 28, SEQ ID No. 29, SEQ ID No. 30, , SEQ ID No. 32, or SEQ ID No. 33 except for one or more amino acid modifications at any one or more of the amino acid residues defined in set 2 or set 4 or set 6 or set 7 (as defined WO2005/066347 and hereinbelow) identified by sequence alignment with SEQ ID No. 34.

[0147] Alternatively the nucleotide sequence encoding a lipid acyltransferase may encode a variant lipid acyltransferase enzyme comprising an amino acid sequence, which amino acid sequence is shown as SEQ ID No. 34, SEQ ID No. 3, SEQ ID No. 4, SEQ ID No. 5, SEQ ID No. 6, SEQ ID No. 7, SEQ ID No. 8, SEQ ID No. 19, SEQ ID No. 10, SEQ ID No. 11, SEQ ID No. 12, SEQ ID No. 13, SEQ ID No. 14, SEQ ID No. 1, SEQ ID No. 15, SEQ ID No. 25, SEQ ID No. 26, SEQ ID No. 27, SEQ ID No. 28, SEQ ID No. 29, SEQ ID No. 30, , SEQ ID No. 32, or SEQ ID No. 33 except for one or more amino acid modifications at any one or more of the amino acid residues defined in set 2 or set 4 or set 6 or set 7 as defined WO2005/066347 and hereinbelow, identified by said parent sequence being structurally aligned with the structural model of P10480 defined herein, which is preferably obtained by structural alignment of P10480 crystal structure coordinates with 1IVN.PDB and/or 1 DEO.PDB as taught within WO2005/066347 and hereinbelow.

[0148] Alternatively, the nucleotide sequence encoding a lipid acyltransferase may encode a variant lipid acyltransferase enzyme comprising an amino acid sequence, which amino acid sequence is shown as SEQ ID No. 34, SEQ ID No. 3, SEQ ID No. 4, SEQ ID No. 5, SEQ ID No. 6, SEQ ID No. 7, SEQ ID No. 8, SEQ ID No. 19, SEQ ID No. 10, SEQ ID No. 11, SEQ ID No. 12, SEQ ID No. 13, SEQ ID No. 14, SEQ ID No. 1, SEQ ID No. 15, SEQ ID No. 25, SEQ ID No. 26, SEQ ID No. 27, SEQ ID No. 28, SEQ ID No. 29, SEQ ID No. 30, SEQ ID No. 32, or SEQ ID No. 33 except for one or more amino acid modifications at any one or more of the amino acid residues taught in set 2 identified when said parent sequence is aligned to the pfam consensus sequence (SEQ ID No. 2) and modified according to a structural model of P10480 to ensure best fit overlap as taught within WO2005/066347 and hereinbelow.

[0149] Preferably, the parent enzyme is an enzyme which comprises, or is homologous to, the amino acid sequence shown as SEQ ID No. 34 and/or SEQ ID No. 15 and/or SEQ ID No. 35.

[0150] Preferably, the nucleotide sequence encoding a lipid acyltransferase may encode a variant enzyme which comprises an amino acid sequence, which amino acid sequence is shown as SEQ ID No. 34 or SEQ ID No. 35 except for one or more amino acid modifications at any one or more of the amino acid residues defined in set 2 or set 4 or set 6 or set 7 as defined in WO2005/066347 and hereinbelow.

DEFINITION OF SETS**Amino acid set 1:**

[0151] Amino acid set 1 (note that these are amino acids in 1IVN - Figure 53 and Figure 54) Gly8, Asp9, Ser10, Leu11, Ser12, Tyr15, Gly44, Asp45, Thr46, Glu69, Leu70, Gly71, Gly72, Asn73, Asp74, Gly75, Lieu76, Gln106, Ile107, Arg108, Leu109, Pro110, Tyr113, Phe121, Phe139, Phe140, Met141, Tyr145, Met151, Asp154, His157, Gly155, Ile156, Pro158

[0152] The highly conserved motifs, such as GDS_x and catalytic residues, were deselected from set 1 (residues underlined). For the avoidance of doubt, set 1 defines the amino acid residues within 10Å of the central carbon atom of a glycerol in the active site of the 1IVN model.

Amino acid set 2:

[0153] Amino acid set 2 (note that the numbering of the amino acids refers to the amino acids in the P10480 mature sequence)

Leu17, Lys22, Met23, Gly40, Asn80, Pro81, Lys82, Asn87, Asn88, Trp111, Val112, Ala114, Tyr117, Leu118, Pro156, Gly159, Gln160, Asn161, Pro162, Ser163, Ala164, Arg165, Ser166, Gln167, Lys168, Val169, Val170, Glu171, Ala172, Tyr179, His180, Asn181, Met209, Leu210, Arg211, Asn215, Lys284, Met285, Gln289 and Val290.

[0154] Table of selected residues in Set 1 compared with Set 2:

IVN model			P10480
IVN	A.hyd homologue		Mature sequence Residue Number
	PFAM	Structure	
Gly8	Gly32		
Asp9	Asp33		
Ser10	Ser34		
Leu11	Leu35		Leu17
Ser12	Ser36		Ser18
			Lys22

IVN model			P10480
IVN	A.hyd homologue		Mature sequence Residue Number
	PFAM	Structure	
			Met23
Tyr15	Gly58		Gly40
Gly44	Asn98		Asn80
Asp45	Pro99		Pro81
Thr46	Lys100		Lys82
			Asn87
			Asn88
Glu69	Trp129		Trp111
Leu70	Val130		Val112
Gly71	Gly131		
Gly72	Ala132		Ala114
Asn73	Asn133		
Asp74	Asp134		
Gly75	Tyr135		Tyr117
Leu76	Leu136		Leu118
Gln106		Pro174	Pro156
Ile107		Gly177	Gly159
Arg108		Gln178	Gln160
Leu109		Asn179	Asn161
Pro110		180 to 190	Pro162
Tyr113			Ser163
			Ala164
			Arg165

IVN model			P10480
IVN	A.hyd homologue		Mature sequence Residue Number
	PFAM	Structure	
			Ser166
			Gln167
			Lys168
			Val169
			Val170
			Glu171
			Ala172
Phe121	His198	Tyr197	Tyr179
		Hls198	Hits180
		Asn199	Asn181
Phe139	Met227		Met209
Phe140	Leu228		Leu210
Melt141	Arg229		Arg211
Tyr145	Asn233		Asn215
			Lys284
Met151	Met303		Met285
Asp154	Asp306		
Gly155	Gln307		Gln289
Ile156	Val308		Val290
His157	His309		
Pro158	Pro310		

Amino acid set 3:

[0155] Amino acid set 3 is identical to set 2 but refers to the *Aeromonas salmonicida* (SEQ ID No. 4) coding sequence, i.e. the amino acid residue numbers are 18 higher in set 3 as this reflects the difference between the amino acid numbering in the mature protein (SEQ ID No. 34) compared with the protein including a signal sequence (SEQ ID No. 25).

5

[0156] The mature proteins of *Aeromonas salmonicida* GDSX (SEQ ID No. 4) and *Aeromonas hydrophila* GDSX (SEQ ID No. 34) differ in five amino acids. These are Thr3Ser, Gln182Lys, Glu309Ala, Ser310Asn, and Gly318-, where the *salmonicida* residue is listed first and the *hydrophila* residue is listed last. The *hydrophila* protein is only 317 amino acids long and lacks a residue in position 318. The *Aeromonas salmonicida* GDSX has considerably high activity on polar lipids such as galactolipid substrates than the *Aeromonas hydrophila* protein. Site scanning was performed on all five amino acid positions.

10

Amino acid set 4:

[0157] Amino acid set 4 is S3, Q182, E309, S310, and -318.

Amino acid set 5:

15 [0158] F13S, D15N, S18G, S18V, Y30F, D116N, D116E, D157 N, Y226F, D228N Y230F.

Amino acid set 6:

[0159] Amino acid set 6 is Ser3, Leu17, Lys22, Met23, Gly40, Asn80, Pro81, Lys82, Asn 87, Asn88, Trp111, Val112, Ala114, Tyr117, Leu118, Pro156, Gly159, Gln160, Asn161, Pro162, Ser163, Ala164, Arg165, Ser166, Gln167, Lys168, Val169, Val170, Glu171, Ala172, Tyr179, His180, Asn181, Gln182, Met209, Leu210, Arg211, Asn215, Lys284, Met285, Gln289, Val290, Glu309, Ser310, -318.

20

[0160] The numbering of the amino acids in set 6 refers to the amino acids residues in P10480 (SEQ ID No. 25) - corresponding amino acids in other sequence backbones can be determined by homology alignment and/or structural alignment to P10480 and/or 1IVN.

25 Amino acid set 7:

[0161] Amino acid set 7 is Ser3, Leu17, Lys22, Met23, Gly40, Asn80, Pro81, Lys82, Asn 87, Asn88, Trp111, Val112, Ala114, Tyr117, Leu118, Pro156, Gly159, Gln160, Asn161, Pro162, Ser163, Ala164, Arg165, Ser166, Gln167, Lys168, Val169, Val170, Glu171, Ala172, Tyr179, His180, Asn181, Gln182, Met209, Leu210, Arg211, Asn215, Lys284, Met285, Gln289, Val290, Glu309, Ser310, -318, Y30X (where X

is selected from A, C, D, E, G, H, I, K, L, M, N, P, Q, R, S, T, V, or W), Y226X (where X is selected from A, C, D, E, G, H, I, K, L, M, N, P, Q, R, S, T, V, or W), Y230X (where X is selected from A, C, D, E, G, H, I, K, L, M, N, P, Q, R, S, T, V, or W), S18X (where X is selected from A, C, D, E, F, H, I, K, L, M, N, P, Q, R, T, W or Y), D157X (where X is selected from A, C, E, F, G, H, I, K, L, M, P, Q, R, S, T, V, W or Y).

5

[0162] The numbering of the amino acids in set 7 refers to the amino acids residues in P10480 (SEQ ID No. 25) - corresponding amino acids in other sequence backbones can be determined by homology alignment and/or structural alignment to P10480 and/or 1IVN).

- 10 **[0163]** The variant enzyme disclosed herein may comprise one or more of the following amino acid modifications compared with the parent enzyme:

S3E, A, G, K, M, Y, R, P, N, T or G

E309Q, R or A, preferably Q or R

-318Y, H, S or Y, preferably Y.

- 15 **[0164]** Preferably, X of the GDSX motif is L. Thus, preferably the parent enzyme comprises the amino acid motif GDSL.

- [0165]** The first parent lipid acyltransferase may comprise any one of the following amino acid sequences: SEQ ID No. 34, SEQ ID No. 3, SEQ ID No. 4, SEQ ID No. 5, SEQ ID No. 6, SEQ ID No. 7, SEQ ID No. 8, SEQ ID No. 19, SEQ ID No. 10, SEQ ID No. 11, SEQ ID No. 12, SEQ ID No. 13, SEQ ID No. 14, SEQ ID No. 1, SEQ ID No. 15, SEQ ID No. 25, SEQ ID No. 26, SEQ ID No. 27, SEQ ID No. 28, SEQ ID No. 29, SEQ ID No. 30, , SEQ ID No. 32 or SEQ ID No. 33.

- 25 **[0166]** Suitably, said second related lipid acyltransferase may comprise any one of the following amino acid sequences: SEQ ID No. 3, SEQ ID No. 34, SEQ ID No. 4, SEQ ID No. 5, SEQ ID No. 6, SEQ ID No. 7, SEQ ID No. 8, SEQ ID No. 19, SEQ ID No. 10, SEQ ID No. 11, SEQ ID No. 12, SEQ ID No. 13, SEQ ID No. 14, SEQ ID No. 1, SEQ ID No. 15, SEQ ID No. 25, SEQ ID No. 26, SEQ ID No. 27, SEQ ID No. 28, SEQ ID No. 29, SEQ ID No. 30, , SEQ ID No. 32 or SEQ ID No. 33.

- 30 **[0167]** The variant enzyme must comprise at least one amino acid modification compared with the parent enzyme. The variant enzyme disclosed herein may comprise at least 2, preferably at least 3, preferably at least

4, preferably at least 5, preferably at least 6, preferably at least 7, preferably at least 8, preferably at least 9, preferably at least 10 amino acid modifications compared with the parent enzyme.

5 **[0168]** When referring to specific amino acid residues herein the numbering is that obtained from alignment of the variant sequence with the reference sequence shown as SEQ ID No. 34 or SEQ ID No. 35.

[0169] The variant enzyme disclosed herein may comprise one or more of the following amino acid substitutions:

10 S3A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, T, V, W, or Y; and/or

L17A, C, D, E, F, G, H, I, K, M, N, P, Q, R, S, T, V, W, or Y; and/or

S18A, C, D, E, F, H, I, K, L, M, N, P, Q, R, T, W, or Y; and/or

K22A, C, D, E, F, G, H, I, L, M, N, P, Q, R, S, T, V, W, or Y; and/or

M23A, C, D, E, F, G, H, I, K, L, N, P, Q, R, S, T, V, W, or Y; and/or

Y30A, C, D, E, G, H, I, K, L, M, N, P, Q, R, S, T, V, or W; and/or

15 G40A, C, D, E, F, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; and/or

N80A, C, D, E, F, G, H, I, K, L, M, P, Q, R, S, T, V, W, or Y; and/or

P81A, C, D, E, F, G, H, I, K, L, M, N, Q, R, S, T, V, W, or Y; and/or

K82A, C, D, E, F, G, H, I, L, M, N, P, Q, R, S, T, V, W, or Y; and/or

N87A, C, D, E, F, G, H, I, K, L, M, P, Q, R, S, T, V, W, or Y; and/or

20 N88A, C, D, E, F, G, H, I, K, L, M, P, Q, R, S, T, V, W, or Y; and/or

W111A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y; and/or

V112A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, W, or Y; and/or

A114C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; and/or

- Y117A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, or W; and/or
- L118A, C, D, E, F, G, H, I, K, M, N, P, Q, R, S, T, V, W, or Y; and/or
- P156A, C, D, E, F, G, H, I, K, L, M, N, Q, R, S, T, V, W, or Y; and/or
- D157A, C, E, F, G, H, I, K, L, M, P, Q, R, S, T, V, W, or Y; and/or
- 5 G159A, C, D, E, F, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; and/or
- Q160A, C, D, E, F, G, H, I, K, L, M, N, P, R, S, T, V, W, or Y; and/or
- N161A, C, D, E, F, G, H, I, K, L, M P, Q, R, S, T, V, W, or Y; and/or
- P162A, C, D, E, F, G, H, I, K, L, M, N, Q, R, S, T, V, W, or Y; and/or
- S163A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, T, V, W, or Y; and/or
- 10 A164C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; and/or
- R165A, C, D, E, F, G, H, I, K, L, M, N, P, Q, S, T, V, W, or Y; and/or
- S166A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, T, V, W, or Y; and/or
- Q167A, C, D, E, F, G, H, I, K, L, M, N, P, R, S, T, V, W, or Y; and/or
- K168A, C, D, E, F, G, H, I, L, M, N, P, Q, R, S, T, V, W, or Y; and/or
- 15 V169A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, W, or Y; and/or
- V170A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, W, or Y; and/or
- E171A, C, D, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; and/or
- A172C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; and/or
- Y179A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, or W; and/or
- 20 H180A, C, D, E, F, G, I, K, L, M, P, Q, R, S, T, V, W, or Y; and/or

N181A, C, D, E, F, G, H, I, K, L, M, P, Q, R, S, T, V, W, or Y; and/or

G182A, C, D, E, F, G, H, I, K, L, M, N, P, R, S, T, V, W, or Y, preferably K; and/or

M209A, C, D, E, F, G, H, I, K, L, N, P, Q, R, S, T, V, W, or Y; and/or

L210 A, C, D, E, F, G, H, I, K, M, N, P, Q, R, S, T, V, W, or Y; and/or

5 R211 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; and/or

N215 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; and/or

Y226A, C, D, E, G, H, I, K, L, M, N, P, Q, R, S, T, V, or W; and/or

Y230A, C, D, E, G, H, I, K, L, M, N, P, Q, R, S, T, V or W; and/or

K284A, C, D, E, F, G, H, I, L, M, N, P, Q, R, S, T, V, W, or Y; and/or

10 M285A, C, D, E, F, G, H, I, K, L, N, P, Q, R, S, T, V, W, or Y; and/or

Q289A, C, D, E, F, G, H, I, K, L, M, N, P, R, S, T, V, W, or Y; and/or

V290A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, W, or Y; and/or

E309A, C, D, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; and/or

S310A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, T, V, W, or Y.

15 **[0170]** In addition or alternatively thereto there may be one or more C-terminal extensions. Preferably the additional C-terminal extension is comprised of one or more aliphatic amino acids, preferably a non-polar amino acid, more preferably of I, L, V or G. Thus, the variant enzyme disclosed herein may comprise one or more of the following C-terminal extensions: 318I, 318L, 318V, 318G.

20 **[0171]** Preferred variant enzymes may have a decreased hydrolytic activity against a phospholipid, such as phosphatidylcholine (PC), may also have an increased transferase activity from a phospholipid.

[0172] Preferred variant enzymes may have an increased transferase activity from a phospholipid, such as phosphatidylcholine (PC), these may also have an increased hydrolytic activity against a phospholipid.

[0173] Modification of one or more of the following residues may result in a variant enzyme having an increased absolute transferase activity against phospholipid:

S3, D157, S310, E309, Y179, N215, K22, Q289, M23, H180, M209, L210, R211, P81, V112, N80, L82, N88; N87

- 5 **[0174]** Specific preferred modifications which may provide a variant enzyme having an improved transferase activity from a phospholipid may be selected from one or more of the following:

S3A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, T, V, W or Y; preferably N, E, K, R, A, P or M, most preferably S3A

10 D157A, C, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y; preferably D157S, R, E, N, G, T, V, Q, K or C

S310A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, T, V, W or Y; preferably S310T -318 E

E309A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, T, V, W or Y; preferably E309 R, E, L, RorA

Y179A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V or W; preferably Y179 D, T, E, R, N, V, K, Q or S, more preferably E, R, N, V, K or Q

15 N215A, C, D, E, F, G, H, I, K, L, M, P, Q, R, S, T, V, W or Y; preferably N215 S, L, R or Y

K22A, C, D, E, F, G, H, I, L, M, N, P, Q, R, S, T, V, W or Y; preferably K22 E, R, C or A

Q289A, C, D, E, F, G, H, I, K, L, M, N, P, R, S, T, V, W or Y; preferably Q289 R, E, G, PorN

M23A, C, D, E, F, G, H, I, K, L, N, P, Q, R, S, T, V, W or Y; preferably M23 K, Q, L, G, TorS

20 H180A, C, D, E, F, G, I, K, L, M, P, Q, R, S, T, V, W or Y; preferably H180 Q, R or K M209 A, C, D, E, F, G, H, I, K, L, N, P, Q, R, S, T, V, W or Y; preferably M209 Q, S, R, A, N, Y, E, V or L

L210A, C, D, E, F, G, H, I, K, M, N, P, Q, R, S, T, V, W or Y; preferably L210 R, A, V, S, T, I, W or M

R211A, C, D, E, F, G, H, I, K, L, M, N, P, Q, S, T, V, W or Y; preferably R211T

P81A, C, D, E, F, G, H, I, K, L, M, N, Q, R, S, T, V, W or Y; preferably P81G

V112A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, W or Y; preferably V112C

N80A, C, D, E, F, G, H, I, K, L, M, P, Q, R, S, T, V, W or Y; preferably N80 R, G, N, D, P, T, E, V, A or G

5 L82A, C, D, E, F, G, H, I, M, N, P, Q, R, S, T, V, W or Y; preferably L82N, S or E N88A, C, D, E, F, G, H, I, K, L, M, P, Q, R, S, T, V, W or Y; preferably N88C

N87A, C, D, E, F, G, H, I, K, L, M, P, Q, R, S, T, V, W or Y; preferably N87M or G

[0175] Preferred modification of one or more of the following residues results in a variant enzyme having an increased absolute transferase activity against phospholipid:

10 S3 N, R, A, G

M23 K, Q, L, G, T, S

H 180 R

L82 G

Y179 E, R, N, V, K or Q

15 E309 R, S, L or A

[0176] One preferred modification is N80D. This is particularly the case when using the reference sequence SEQ ID No. 35 as the backbone. Thus, the reference sequence may be SEQ ID No. 16. This modification may be in combination with one or more further modifications. Therefore the nucleotide sequence encoding a lipid acyltransferase for use in any one of the host cells, vectors, methods and uses disclosed herein may encode a
20 lipid acyltransferase that comprises SEQ ID No. 35.

[0177] As noted above, when referring to specific amino acid residues herein the numbering is that obtained from alignment of the variant sequence with the reference sequence shown as SEQ ID No. 34 or SEQ ID No.
35

25

[0178] Much by preference, the nucleotide sequence encoding a lipid acyltransferase for use in any one of the host cells, vectors, methods and uses disclosed herein may encode a lipid comprising the amino acid sequence

shown as SEQ ID No. 16, or an amino acid sequence which has 75% or more, preferably 85% or more, more preferably 90% or more, even more preferably 95% or more, even more preferably 98% or more, or even more preferably 99% or more identity to SEQ ID No. 16. This enzyme may be considered a variant enzyme.

- 5 **[0179]** For the purposes disclosed herein, the degree of identity is based on the number of sequence elements which are the same. The degree of identity for amino acid sequences may be suitably determined by means of computer programs known in the art, such as Vector NTI 10 (Invitrogen Corp.). For pairwise alignment the score used is preferably BLOSUM62 with Gap opening penalty of 10.0 and Gap extension penalty of 0.1.
- 10 **[0180]** Suitably, the degree of identity with regard to an amino acid sequence is determined over at least 20 contiguous amino acids, preferably over at least 30 contiguous amino acids, preferably over at least 40 contiguous amino acids, preferably over at least 50 contiguous amino acids, preferably over at least 60 contiguous amino acids.
- 15 **[0181]** Suitably, the degree of identity with regard to an amino acid sequence may be determined over the whole sequence.
- 20 **[0182]** Suitably, the nucleotide sequence encoding a lipid acyltransferase/ lipid acyl transferase enzyme disclosed herein may be obtainable, preferably obtained, from organisms from one or more of the following genera: *Aeromonas*, *Streptomyces*, *Saccharomyces*, *Lactococcus*, *Mycobacterium*, *Streptococcus*, *Lactobacillus*, *Desulfotobacterium*, *Bacillus*, *Campylobacter*, *Vibrionaceae*, *Xylella*, *Sulfolobus*, *Aspergillus*, *Schizosaccharomyces*, *Listeria*, *Neisseria*, *Mesorhizobium*, *Ralstonia*, *Xanthomonas*, *Candida*, *Thermobifida* and *Corynebacterium*.
- 25 **[0183]** Suitably, the nucleotide sequence encoding a lipid acyltransferase / lipid acyl transferase enzyme disclosed herein may be obtainable, preferably obtained, from one or more of the following organisms:
- 30 *Aeromonas hydrophila*, *Aeromonas salmonicida*, *Streptomyces coelicolor*, *Streptomyces rimosus*, *Mycobacterium*, *Streptococcus pyogenes*, *Lactococcus lactis*, *Streptococcus pyogenes*, *Streptococcus thermophilus*, *Streptomyces thermosacchari*, *Streptomyces avermitilis* *Lactobacillus helveticus*, *Desulfotobacterium dehalogenans*, *Bacillus sp*, *Campylobacter jejuni*, *Vibrionaceae*, *Xylella fastidiosa*, *Sulfolobus solfataricus*, *Saccharomyces cerevisiae*, *Aspergillus terreus*, *Schizosaccharomyces pombe*, *Listeria innocua*, *Listeria monocytogenes*, *Neisseria meningitidis*, *Mesorhizobium loti*, *Ralstonia solanacearum*, *Xanthomonas campestris*, *Xanthomonas axonopodis* *Candida parapsilosis* *Thermobifida fusca* and *Corynebacterium efficiens*.

[0184] In one aspect, preferably the nucleotide sequence encoding a lipid acyltransferase for use in any methods of the present invention encodes a lipid acyl transferase enzyme according to the present invention obtainable, preferably obtained or derived, from *Aeromonas spp.*, or *Aeromonas salmonicida*.

5 [0185] Enzymes which function as lipid acyltransferases in accordance with the present invention can be routinely identified using the assay taught in Example 12 of VO2004/064537. Using this assay, in which there is a very high water content - approximately 95%, lipid acyltransferases/lipid acyl transferase in accordance with the present invention are those which have at least 2% acyltransferase activity (relative transferase activity), preferably at least 5% relative transferase activity, preferably at least 10% relative transferase
10 activity, preferably at least 15%, 20%, 25% 26%, 28%, 30%, 40% 50%, 60% or 75% relative transferase activity.

[0186] Phospholipases may act as acyl-transferase enzymes in low water environments. Therefore it is considered that in place of or in addition to the phospholipid acyltransferase enzyme a phospholipase enzyme
15 may be used when process for the modification of the edible oil of fat takes place in a low water environment.

[0187] The term "high water" as used herein means any substrate or foodstuff with more than 3% water content, preferably more than 4%, more than 5%, more than 6%, more than 7%, more than 8%, more than 9%, more than 10%, more than 20%, more than 30%, more than 40%, more than 50%, more than 60%, more than
20 70%, more than 80% or more than 90%.

[0188] The term "low water" as used herein means any substrate or foodstuff with less than 3% water content, preferably less than 2%, less than 1% or less than 0.5%, less than 0.3%, less than 0.2, less than 0.1, less than 0.05, or less than 0.01 %

25 [0189] For avoidance of doubt milk is a high water environment where as butterfat is a low water environment.

[0190] Suitable phospholipases for uses disclosed herein include phospholipase A1, phospholipase A2, or phospholipase B. Phospholipase A1, phospholipase A2, or phospholipase B may also be used in co-ordination
30 with the lipid acyl transferase activity. Phospholipase C and /or D may also be used in co-ordination with the lipid acyl transferase activity/phospholipase A1, A2 and/or B activity in analogy with WO2005/089562. Preferred phospholipases may include phospholipase A2, such as LecitaseTM or the *Fusarium venenatum* and *Tuber albidum* phospholipase disclosed in WO2004/97012 (Novozymes/Chr. Hansen). A *Fusarium*
35 *venenatum* phospholipase is sold by Novozymes as MAX YIELDTM.

ISOLATED

[0191] In one aspect, the method of the present invention comprises the additional step of recovering/isolating the lipid acyltransferase. Thus, the lipid acyltransferase produced may be in an isolated form.

5

[0192] In another aspect, the nucleotide sequence encoding a lipid acyltransferase for use in the present invention may be in an isolated form.

[0193] The term "isolated" means that the sequence or protein is at least substantially free from at least one other component with which the sequence or protein is naturally associated in nature and as found in nature.

10

PURIFIED

[0194] In one aspect, the method of the present invention comprises the additional step of purifying the lipid acyltransferase.

15

[0195] In another aspect, the nucleotide sequence encoding a lipid acyltransferase for use in the present invention may be in a purified form.

[0196] The term "purified" means that the sequence is in a relatively pure state - e.g. at least about 51% pure, or at least about 75%, or at least about 80%, or at least about 90% pure, or at least about 95% pure or at least about 98% pure.

20

CLONING A NUCLEOTIDE SEQUENCE ENCODING A POLYPEPTIDE ACCORDING TO THE PRESENT INVENTION

[0197] A nucleotide sequence encoding either a polypeptide which has the specific properties as defined herein or a polypeptide which is suitable for modification may be isolated from any cell or organism producing said polypeptide. Various methods are well known within the art for the isolation of nucleotide sequences.

25

[0198] For example, a genomic DNA and/or cDNA library may be constructed using chromosomal DNA or messenger RNA from the organism producing the polypeptide. If the amino acid sequence of the polypeptide is known, labeled oligonucleotide probes may be synthesised and used to identify polypeptide-encoding clones from the genomic library prepared from the organism. Alternatively, a labelled oligonucleotide probe

30

containing sequences homologous to another known polypeptide gene could be used to identify polypeptide-encoding clones. In the latter case, hybridisation and washing conditions of lower stringency are used.

[0199] Alternatively, polypeptide-encoding clones could be identified by inserting fragments of genomic

5 DNA into an expression vector, such as a plasmid, transforming enzyme-negative bacteria with the resulting genomic DNA library, and then plating the transformed bacteria onto agar containing an enzyme inhibited by the polypeptide, thereby allowing clones expressing the polypeptide to be identified.

[0200] In a yet further alternative, the nucleotide sequence encoding the polypeptide may be prepared

10 synthetically by established standard methods, e.g. the phosphoroamidite method described by Beaucage S.L. et al (1981) Tetrahedron Letters 22, p 1859-1869, or the method described by Matthes et al (1984) EMBO J. 3, p 801-805. In the phosphoroamidite method, oligonucleotides are synthesised, e.g. in an automatic DNA synthesiser, purified, annealed, ligated and cloned in appropriate vectors.

15 [0201] The nucleotide sequence may be of mixed genomic and synthetic origin, mixed synthetic and cDNA origin, or mixed genomic and cDNA origin, prepared by ligating fragments of synthetic, genomic or cDNA origin (as appropriate) in accordance with standard techniques. Each ligated fragment corresponds to various parts of the entire nucleotide sequence. The DNA sequence may also be prepared by polymerase chain

20 reaction (PCR) using specific primers, for instance as described in US 4,683,202 or in Saiki R K et al (Science (1988) 239, pp 487-4991).

NUCLEOTIDE SEQUENCES

[0202] The present invention also encompasses nucleotide sequences encoding polypeptides having the specific properties as defined herein. The term "nucleotide sequence" as used herein refers to an

25 oligonucleotide sequence or polynucleotide sequence, and variant, homologues, fragments and derivatives thereof (such as portions thereof). The nucleotide sequence may be of genomic or synthetic or recombinant origin, which may be double-stranded or single-stranded whether representing the sense or antisense strand.

[0203] The term "nucleotide sequence" in relation to the present invention includes genomic DNA, cDNA, synthetic DNA, and RNA. Preferably it means DNA, more preferably cDNA for the coding sequence.

30

[0204] In a preferred embodiment, the nucleotide sequence *per se* encoding a polypeptide having the specific properties as defined herein does not cover the native nucleotide sequence in its natural environment when it is linked to its naturally associated sequence(s) that is/are also in its/their natural environment. For ease of reference, we shall call this preferred embodiment the "non-native nucleotide sequence". In this regard, the

term "native nucleotide sequence" means an entire nucleotide sequence that is in its native environment and when operatively linked to an entire promoter with which it is naturally associated, which promoter is also in its native environment. Thus, the polypeptide of the present invention can be expressed by a nucleotide sequence in its native organism but wherein the nucleotides sequence is not under the control of the promoter with which it is naturally associated within that organism.

[0205] Preferably the polypeptide is not a native polypeptide. In this regard, the term "native polypeptide" means an entire polypeptide that is in its native environment and when it has been expressed by its native nucleotide sequence.

[0206] Typically, the nucleotide sequence encoding polypeptides having the specific properties as defined herein is prepared using recombinant DNA techniques (i.e. recombinant DNA). However, in an alternative embodiment of the invention, the nucleotide sequence could be synthesised, in whole or in part, using chemical methods well known in the art (see Caruthers MH et al (1980) Nuc Acids Res Symp Ser 215-23 and Horn T et al (1980) Nuc Acids Res Symp Ser 225-232).

MOLECULAR EVOLUTION

[0207] Once an enzyme-encoding nucleotide sequence has been isolated, or a putative enzyme-encoding nucleotide sequence has been identified, it may be desirable to modify the selected nucleotides sequence, for example it may be desirable to mutate the sequence in order to prepare an enzyme in accordance with the present invention.

[0208] Mutations may be introduced using synthetic oligonucleotides. These oligonucleotides contain nucleotide sequences flanking the desired mutation sites.

[0209] A suitable method is disclosed in Morinaga et al (Biotechnology (1984) 2, p646-649). Another method of introducing mutations into enzyme-encoding nucleotide sequences is described in Nelson and Long (Analytical Biochemistry (1989), 180, p 147-151).

[0210] Instead of site directed mutagenesis, such as described above, one can introduce mutations randomly for instance using a commercial kit such as the GeneMorph PCR mutagenesis kit from Stratagene, or the Diversify PCR random mutagenesis kit from Clontech. EP 0 583 265 refers to methods of optimising PCR based mutagenesis, which can also be combined with the use of mutagenic DNA analogues such as those described in EP 0 866 796. Error prone PCR technologies are suitable for the production of variants of lipid acyl transferases with preferred characteristics. WO0206457 refers to molecular evolution of lipases.

[0211] A third method to obtain novel sequences is to fragment non-identical nucleotide sequences, either by using any number of restriction enzymes or an enzyme such as Dnase I, and reassembling full nucleotide sequences coding for functional proteins. Alternatively one can use one or multiple non-identical nucleotide sequences and introduce mutations during the reassembly of the full nucleotide sequence. DNA shuffling and family shuffling technologies are suitable for the production of variants of lipid acyl transferases with preferred characteristics. Suitable methods for performing 'shuffling' can be found in EP0 752 008, EP1 138 763, EP1 103 606. Shuffling can also be combined with other forms of DNA mutagenesis as described in US 6,180,406 and WO 01/34835.

[0212] Thus, it is possible to produce numerous site directed or random mutations into a nucleotide sequence, either *in vivo* or *in vitro*, and to subsequently screen for improved functionality of the encoded polypeptide by various means. Using *in silico* and *ex vivo* mediated recombination methods (see WO 00/58517, US 6,344,328, US 6,361,974), for example, molecular evolution can be performed where the variant produced retains very low homology to known enzymes or proteins. Such variants thereby obtained may have significant structural analogy to known transferase enzymes, but have very low amino acid sequence homology.

[0213] As a non-limiting example, In addition, mutations or natural variants of a polynucleotide sequence can be recombined with either the wild type or other mutations or natural variants to produce new variants. Such new variants can also be screened for improved functionality of the encoded polypeptide.

[0214] The application of the above-mentioned and similar molecular evolution methods allows the identification and selection of variants of the enzymes of the present invention which have preferred characteristics without any prior knowledge of protein structure or function, and allows the production of non-predictable but beneficial mutations or variants. There are numerous examples of the application of molecular evolution in the art for the optimisation or alteration of enzyme activity, such examples include, but are not limited to one or more of the following: optimised expression and/or activity in a host cell or *in vitro*, increased enzymatic activity, altered substrate and/or product specificity, increased or decreased enzymatic or structural stability, altered enzymatic activity/specificity in preferred environmental conditions, e.g. temperature, pH, substrate

[0215] As will be apparent to a person skilled in the art, using molecular evolution tools an enzyme may be altered to improve the functionality of the enzyme.

[0216] Suitably, the nucleotide sequence encoding a lipid acyltransferase used in the disclosure may encode a variant lipid acyltransferase, i.e. the lipid acyltransferase may contain at least one amino acid substitution,

deletion or addition, when compared to a parental enzyme. Variant enzymes retain at least 1%, 2%, 3%, 5%, 10%, 15%, 20%, 30%, 40%, 50 %, 60%, 70%, 80%, 90%, 95%, 97%, 99% homology with the parent enzyme. Suitable parent enzymes may include any enzyme with esterase or lipase activity. Preferably, the parent enzyme aligns to the pfam00657 consensus sequence.

5

[0217] In a preferable embodiment a variant lipid acyltransferase enzyme retains or incorporates at least one or more of the pfam00657 consensus sequence amino acid residues found in the GDSx, GANDY and HPT blocks.

10 **[0218]** Enzymes, such as lipases with no or low lipid acyltransferase activity in an aqueous environment may be mutated using molecular evolution tools to introduce or enhance the transferase activity, thereby producing a lipid acyltransferase enzyme with significant transferase activity suitable for use in the compositions and methods of the present invention.

15 **[0219]** Suitably, the nucleotide sequence encoding a lipid acyltransferase for use in any one of the vectors, host cells, methods and/or uses of the present disclosure may encode a lipid acyltransferase that may be a variant with enhanced enzyme activity on polar lipids, preferably phospholipids and/or glycolipids when compared to the parent enzyme. Preferably, such variants also have low or no activity on lyso polar lipids. The enhanced activity on polar lipids, phospholipids and/or glycolipids may be the result of hydrolysis and/or
20 transferase activity or a combination of both,

[0220] Variant lipid acyltransferases may have decreased activity on triglycerides, and/or monoglycerides and/or diglycerides compared with the parent enzyme.

25 **[0221]** Suitably the variant enzyme may have no activity on triglycerides and/or monoglycerides and/or diglycerides.

[0222] Alternatively, the variant enzyme may have increased activity on triglycerides, and/or may also have increased activity on one or more of the following, polar lipids, phospholipids, lecithin, phosphatidylcholine, glycolipids, digalactosyl monoglyceride, monogalactosyl monoglyceride.
30

[0223] Variants of lipid acyltransferases are known, and one or more of such variants may be suitable for use in the methods and uses according to the present disclosure and/or in the enzyme compositions according to the present disclosure. By way of example only, variants of lipid acyltransferases are described in the
35 following references may be used in accordance with the present disclosure: Hilton & Buckley J Biol. Chem.

1991 Jan 15; 266 (2): 997-1000; Robertson et al J. Biol. Chem. 1994 Jan 21; 269(3):2146-50; Brumlik et al J. Bacteriol 1996 Apr; 178 (7): 2060-4; Peelman et al Protein Sci. 1998 Mar; 7(3):587-99.

AMINO ACID SEQUENCES

5 [0224] The present invention also encompasses amino acid sequences encoded by a nucleotide sequence which encodes a lipid acyltransferase for use in any one of the vectors, host cells, methods and/or uses of the present invention.

[0225] As used herein, the term "amino acid sequence" is synonymous with the term "polypeptide" and/or the
10 term "protein". In some instances, the term "amino acid sequence" is synonymous with the term "peptide".

[0226] The amino acid sequence may be prepared/isolated from a suitable source, or it may be made synthetically or it may be prepared by use of recombinant DNA techniques.

15 [0227] Suitably, the amino acid sequences may be obtained from the isolated polypeptides taught herein by standard techniques.

[0228] One suitable method for determining amino acid sequences from isolated polypeptides is as follows:

20 [0229] Purified polypeptide may be freeze-dried and 100 µg of the freeze-dried material may be dissolved in 50 µl of a mixture of 8 M urea and 0.4 M ammonium hydrogen carbonate, pH 8.4. The dissolved protein may be denatured and reduced for 15 minutes at 50°C following overlay with nitrogen and addition of 5 µl of 45 mM dithiothreitol. After cooling to room temperature, 5 µl of 100 mM iodoacetamide may be added for the cysteine residues to be derivatized for 15 minutes at room temperature in the dark under nitrogen.

25 [0230] 135 µl of water and 5 µg of endoproteinase Lys-C in 5 µl of water may be added to the above reaction mixture and the digestion may be carried out at 37°C under nitrogen for 24 hours.

[0231] The resulting peptides may be separated by reverse phase HPLC on a VYDAC C18 column
30 (0.46x15cm;10µm; The Separation Group, California, USA) using solvent A: 0.1% TFA in water and solvent B: 0.1% TFA in acetonitrile. Selected peptides may be re-chromatographed on a Develosil C18 column using the same solvent system, prior to N-terminal sequencing. Sequencing may be done using an Applied Biosystems 476A sequencer using pulsed liquid fast cycles according to the manufacturers instructions (Applied Biosystems, California, USA).

SEQUENCE IDENTITY OR SEQUENCE HOMOLOGY

[0232] Here, the term "homologue" means an entity having a certain homology with the subject amino acid sequences and the subject nucleotide sequences. Here, the term "homology" can be equated with "identity".

5 [0233] The homologous amino acid sequence and/or nucleotide sequence should provide and/or encode a polypeptide which retains the functional activity and/or enhances the activity of the enzyme.

[0234] In the present context, a homologous sequence is taken to include an amino acid sequence which may be at least 75, 85 or 90% identical, preferably at least 95 or 98% identical to the subject sequence. Typically,
10 the homologues will comprise the same active sites etc. as the subject amino acid sequence. Although homology can also be considered in terms of similarity (i.e. amino acid residues having similar chemical properties/functions), in the context of the present invention it is preferred to express homology in terms of sequence identity.

15 [0235] In the present context, a homologous sequence is taken to include a nucleotide sequence which may be at least 75, 85 or 90% identical, preferably at least 95 or 98% identical to a nucleotide sequence encoding a polypeptide of the present invention (the subject sequence). Typically, the homologues will comprise the same sequences that code for the active sites etc. as the subject sequence. Although homology can also be considered in terms of similarity (i.e. amino acid residues having similar chemical properties/functions), in the
20 context of the present invention it is preferred to express homology in terms of sequence identity.

[0236] Homology comparisons can be conducted by eye, or more usually, with the aid of readily available sequence comparison programs. These commercially available computer programs can calculate % homology between two or more sequences.

25 [0237] % homology may be calculated over contiguous sequences, i.e. one sequence is aligned with the other sequence and each amino acid in one sequence is directly compared with the corresponding amino acid in the other sequence, one residue at a time. This is called an "ungapped" alignment. Typically, such ungapped alignments are performed only over a relatively short number of residues.

30 [0238] Although this is a very simple and consistent method, it fails to take into consideration that, for example, in an otherwise identical pair of sequences, one insertion or deletion will cause the following amino acid residues to be put out of alignment, thus potentially resulting in a large reduction in % homology when a global alignment is performed. Consequently, most sequence comparison methods are designed to produce

optimal alignments that take into consideration possible insertions and deletions without penalising unduly the overall homology score. This is achieved by inserting "gaps" in the sequence alignment to try to maximise local homology. However, these more complex methods assign "gap penalties" to each gap that occurs in the alignment so that, for the same number of identical amino acids, a sequence alignment with as few gaps as possible - reflecting higher relatedness between the two compared sequences - will achieve a higher score than one with many gaps. "Affine gap costs" are typically used that charge a relatively high cost for the existence of a gap and a smaller penalty for each subsequent residue in the gap. This is the most commonly used gap scoring system. High gap penalties will of course produce optimised alignments with fewer gaps. Most alignment programs allow the gap penalties to be modified. However, it is preferred to use the default values when using such software for sequence comparisons.

Calculation of maximum % homology therefore firstly requires the production of an optimal alignment, taking into consideration gap penalties. A suitable computer program for carrying out such an alignment is the Vector NTI (Invitrogen Corp.). Examples of other software that can perform sequence comparisons include, but are not limited to, the BLAST package (see Ausubel et al 1999 Short Protocols in Molecular Biology, 4th Ed - Chapter 18), and FASTA (Altschul et al 1990 J. Mol. Biol. 403-410). Both BLAST and FASTA are available for offline and online searching (see Ausubel *et al* 1999, pages 7-58 to 7-60). However, for some applications, it is preferred to use the Vector NTI program. A new tool, called BLAST 2 Sequences is also available for comparing protein and nucleotide sequence (see FEMS Microbiol Lett 1999 174(2): 247-50; FEMS Microbiol Lett 1999 177(1): 187-8 and tatiana@nobi.nlm.nih.gov).

[0239] Although the final % homology can be measured in terms of identity, the alignment process itself is typically not based on an all-or-nothing pair comparison. Instead, a scaled similarity score matrix is generally used that assigns scores to each pairwise comparison based on chemical similarity or evolutionary distance. An example of such a matrix commonly used is the BLOSUM62 matrix - the default matrix for the BLAST suite of programs. Vector NTI programs generally use either the public default values or a custom symbol comparison table if supplied (see user manual for further details). For some applications, it is preferred to use the default values for the Vector NTI package.

[0240] Alternatively, percentage homologies may be calculated using the multiple alignment feature in Vector NTI (Invitrogen Corp.), based on an algorithm, analogous to CLUSTAL (Higgins DG & Sharp PM (1988), Gene 73(1), 237-244).

[0241] Once the software has produced an optimal alignment, it is possible to calculate % homology, preferably % sequence identity. The software typically does this as part of the sequence comparison and generates a numerical result.

[0242] Should Gap Penalties be used when determining sequence identity, then preferably the following parameters are used for pairwise alignment:

FOR BLAST			
GAP OPEN			0
GAP EXTENSION			0
FOR CLUSTAL	DNA	PROTEIN	
WORD SIZE	2	1	K triple
GAP PENALTY	15	10	
GAP EXTENSION	6.66	0.1	

- 5 [0243] In one embodiment, preferably the sequence identity for the nucleotide sequences is determined using CLUSTAL with the gap penalty and gap extension set as defined above.

[0244] Suitably, the degree of identity with regard to a nucleotide sequence is determined over at least 20 contiguous nucleotides, preferably over at least 30 contiguous nucleotides, preferably over at least 40
 10 contiguous nucleotides, preferably over at least 50 contiguous nucleotides, preferably over at least 60 contiguous nucleotides, preferably over at least 100 contiguous nucleotides.

[0245] Suitably, the degree of identity with regard to a nucleotide sequence may be determined over the
 15 whole sequence.

[0246] In one embodiment the degree of amino acid sequence identity in accordance with the present invention may be suitably determined by means of computer programs known in the art, such as Vector NTI 10 (Invitrogen Corp.). For pairwise alignment the matrix used is preferably BLOSUM62 with Gap opening
 20 penalty of 10.0 and Gap extension penalty of 0.1.

[0247] Suitably, the degree of identity with regard to an amino acid sequence is determined over at least 20 contiguous amino acids, preferably over at least 30 contiguous amino acids, preferably over at least 40
 25 contiguous amino acids, preferably over at least 50 contiguous amino acids, preferably over at least 60 contiguous amino acids.

[0248] Suitably, the degree of identity with regard to an amino acid sequence may be determined over the whole sequence.

[0249] The sequences may also have deletions, insertions or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent substance. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues as long as the secondary binding activity of the substance is retained. For example, negatively charged amino acids include aspartic acid and glutamic acid; positively charged amino acids include lysine and arginine; and amino acids with uncharged polar head groups having similar hydrophilicity values include leucine, isoleucine, valine, glycine, alanine, asparagine, glutamine, serine, threonine, phenylalanine, and tyrosine.

[0250] Conservative substitutions may be made, for example according to the Table below. Amino acids in the same block in the second column and preferably in the same line in the third column may be substituted for each other:

ALIPHATIC	Non-polar	G A P
		I L V
	Polar - uncharged	C S T M
		N Q
	Polar - charged	D E
		K R
AROMATIC		H F W Y

[0251] The present disclosure also encompasses homologous substitution (substitution and replacement are both used herein to mean the interchange of an existing amino acid residue, with an alternative residue) that may occur i.e. like-for-like substitution such as basic for basic, acidic for acidic, polar for polar etc. Non-homologous substitution may also occur i.e. from one class of residue to another or alternatively involving the inclusion of unnatural amino acids such as ornithine (hereinafter referred to as Z), diaminobutyric acid ornithine (hereinafter referred to as B), norleucine ornithine (hereinafter referred to as O), pyriylalanine, thienylalanine, naphthylalanine and phenylglycine.

[0252] Replacements may also be made by unnatural amino acids.

[0253] Variant amino acid sequences may include suitable spacer groups that may be inserted between any two amino acid residues of the sequence including alkyl groups such as methyl, ethyl or propyl groups in addition to amino acid spacers such as glycine or β -alanine residues. A further form of variation, involves the presence of one or more amino acid residues in peptoid form, will be well understood by those skilled in the art. For the avoidance of doubt, "the peptoid form" is used to refer to variant amino acid residues wherein the α -carbon substituent group is on the residue's nitrogen atom rather than the α -carbon. Processes for preparing peptides in the peptoid form are known in the art, for example Simon RJ et al., PNAS (1992) 89(20), 9367-9371 and Horwell DC, Trends Biotechnol. (1995) 13(4), 132-134.

[0254] Nucleotide sequences for use in the present invention or encoding a polypeptide having the specific properties defined herein may include within them synthetic or modified nucleotides. A number of different types of modification to oligonucleotides are known in the art. These include methylphosphonate and phosphorothioate backbones and/or the addition of acridine or polylysine chains at the 3' and/or 5' ends of the molecule. For the purposes of the present invention, it is to be understood that the nucleotide sequences described herein may be modified by any method available in the art. Such modifications may be carried out in order to enhance the *in vivo* activity or life span of nucleotide sequences.

[0255] The present disclosure also encompasses the use of nucleotide sequences that are complementary to the sequences discussed herein, or any derivative, fragment or derivative thereof. If the sequence is complementary to a fragment thereof then that sequence can be used as a probe to identify similar coding sequences in other organisms etc.

[0256] Polynucleotides which are not 100% homologous to the sequences of the present invention but fall within the scope of the invention can be obtained in a number of ways. Other variants of the sequences described herein may be obtained for example by probing DNA libraries made from a range of individuals, for example individuals from different populations. In addition, other viral/bacterial, or cellular homologues particularly cellular homologues found in mammalian cells (e.g. rat, mouse, bovine and primate cells), may be obtained and such homologues and fragments thereof in general will be capable of selectively hybridising to the sequences shown in the sequence listing herein. Such sequences may be obtained by probing cDNA libraries made from or genomic DNA libraries from other animal species, and probing such libraries with probes comprising all or part of any one of the sequences in the attached sequence listings under conditions of medium to high stringency. Similar considerations apply to obtaining species homologues and allelic variants of the polypeptide or nucleotide sequences of the invention.

[0257] Variants and strain/species homologues may also be obtained using degenerate PCR which will use

primers designed to target sequences within the variants and homologues encoding conserved amino acid sequences within the sequences of the present invention. Conserved sequences can be predicted, for example, by aligning the amino acid sequences from several variants/homologues. Sequence alignments can be performed using computer software known in the art. For example the GCG Wisconsin PileUp program is widely used.

[0258] The primers used in degenerate PCR will contain one or more degenerate positions and will be used at stringency conditions lower than those used for cloning sequences with single sequence primers against known sequences.

[0259] Alternatively, such polynucleotides may be obtained by site directed mutagenesis of characterised sequences. This may be useful where for example silent codon sequence changes are required to optimise codon preferences for a particular host cell in which the polynucleotide sequences are being expressed. Other sequence changes may be desired in order to introduce restriction polypeptide recognition sites, or to alter the property or function of the polypeptides encoded by the polynucleotides.

[0260] Polynucleotides (nucleotide sequences) of the invention may be used to produce a primer, e.g. a PCR primer, a primer for an alternative amplification reaction, a probe e.g. labelled with a revealing label by conventional means using radioactive or nonradioactive labels, or the polynucleotide may be cloned into vectors. Such primers, probes and other fragments will be at least 15, preferably at least 20, for example at least 25, 30 or 40 nucleotides in length, and are also encompassed by the term polynucleotides of the disclosure as used herein.

[0261] Polynucleotides such as DNA polynucleotides and probes according to the invention may be produced recombinantly, synthetically, or by any means available to those of skill in the art. They may also be cloned by standard techniques.

[0262] In general, primers will be produced by synthetic means, involving a stepwise manufacture of the desired nucleic acid sequence one nucleotide at a time. Techniques for accomplishing this using automated techniques are readily available in the art.

[0263] Longer polynucleotides will generally be produced using recombinant means, for example using a PCR (polymerase chain reaction) cloning techniques. This will involve making a pair of primers (e.g. of about 15 to 30 nucleotides) flanking a region of the lipid targeting sequence which it is desired to clone, bringing the primers into contact with mRNA or cDNA obtained from an animal or human cell, performing a polymerase

chain reaction under conditions which bring about amplification of the desired region, isolating the amplified fragment (e.g. by purifying the reaction mixture on an agarose gel) and recovering the amplified DNA. The primers may be designed to contain suitable restriction enzyme recognition sites so that the amplified DNA can be cloned into a suitable cloning vector.

5 HYBRIDISATION

[0264] The present disclosure also encompasses sequences that are complementary to the sequences of the present invention or sequences that are capable of hybridising either to the sequences of the present invention or to sequences that are complementary thereto.

10 [0265] The term "hybridisation" as used herein shall include "the process by which a strand of nucleic acid joins with a complementary strand through base pairing" as well as the process of amplification as carried out in polymerase chain reaction (PCR) technologies.

[0266] The present disclosure also encompasses the use of nucleotide sequences that are capable of
15 hybridising to the sequences that are complementary to the subject sequences discussed herein, or any derivative, fragment or derivative thereof.

[0267] The present disclosure also encompasses sequences that are complementary to sequences that are capable of hybridising to the nucleotide sequences discussed herein.

20

[0268] Hybridisation conditions are based on the melting temperature (T_m) of the nucleotide binding complex, as taught in Berger and Kimmel (1987, Guide to Molecular Cloning Techniques, Methods in Enzymology, Vol. 152, Academic Press, San Diego CA), and confer a defined "stringency" as explained below.

25

[0269] Maximum stringency typically occurs at about $T_m - 5^\circ\text{C}$ (5°C below the T_m of the probe); high stringency at about 5°C to 10°C below T_m ; intermediate stringency at about 10°C to 20°C below T_m ; and low stringency at about 20°C to 25°C below T_m . As will be understood by those of skill in the art, a maximum stringency hybridisation can be used to identify or detect identical nucleotide sequences while an intermediate
30 (or low) stringency hybridisation can be used to identify or detect similar or related polynucleotide sequences.

[0270] Preferably, the present disclosure encompasses sequences that are complementary to sequences that are capable of hybridising under high stringency conditions or intermediate stringency conditions to nucleotide sequences encoding polypeptides having the specific properties as defined herein.

[0271] More preferably, the present disclosure encompasses sequences that are complementary to sequences that are capable of hybridising under high stringent conditions (e.g. 65°C and 0,1xSSC {1xSSC = 0.15 M NaCl, 0.015 M Na-citrate pH 7.0}) to nucleotide sequences encoding polypeptides having the specific properties as defined herein.

5

[0272] The present disclosure also relates to nucleotide sequences that can hybridise to the nucleotide sequences discussed herein (including complementary sequences of those discussed herein).

10

[0273] The present disclosure also relates to nucleotide sequences that are complementary to sequences that can hybridise to the nucleotide sequences discussed herein (including complementary sequences of those discussed herein).

15

[0274] Also included within the scope of the present disclosure are polynucleotide sequences that are capable of hybridising to the nucleotide sequences discussed herein under conditions of intermediate to maximal stringency.

20

[0275] In a preferred aspect, the present disclosure covers nucleotide sequences that can hybridise to the nucleotide sequences discussed herein, or the complement thereof, under stringent conditions (e.g. 50°C and 0.2xSSC).

[0276] In a more preferred aspect, the present disclosure covers nucleotide sequences that can hybridise to the nucleotide sequences discussed herein, or the complement thereof, under high stringent conditions (e.g. 65°C and 0.1xSSC).

EXPRESSION OF POLYPEPTIDES

25

[0277] A nucleotide sequence for use in the present invention or for encoding a polypeptide having the specific properties as defined herein can be incorporated into a recombinant replicable vector. The vector may be used to replicate and express the nucleotide sequence, in polypeptide form, in and/or from a compatible host cell. Expression may be controlled using control sequences which include promoters/enhancers and other expression regulation signals. Prokaryotic promoters and promoters functional in eukaryotic cells may be used. Tissue specific or stimuli specific promoters may be used. Chimeric promoters may also be used comprising sequence elements from two or more different promoters described above.

30

[0278] The polypeptide produced by a host recombinant cell by expression of the nucleotide sequence may be secreted or may be contained intracellularly depending on the sequence and/or the vector used. The coding

sequences can be designed with signal sequences which direct secretion of the substance coding sequences through a particular prokaryotic or eukaryotic cell membrane.

CONSTRUCTS

[0279] The term "construct" - which is synonymous with terms such as "conjugate", "cassette" and "hybrid" - includes a nucleotide sequence encoding a polypeptide having the specific properties as defined herein for use according to the present invention directly or indirectly attached to a promoter. An example of an indirect attachment is the provision of a suitable spacer group such as an intron sequence, such as the Sh1-intron or the ADH intron, intermediate the promoter and the nucleotide sequence of the present invention. The same is true for the term "fused" in relation to the present invention which includes direct or indirect attachment. In some cases, the terms do not cover the natural combination of the nucleotide sequence coding for the protein ordinarily associated with the wild type gene promoter and when they are both in their natural environment.

[0280] The construct may even contain or express a marker which allows for the selection of the genetic construct.

[0281] For some applications, preferably the construct comprises at least a nucleotide sequence of the present invention or a nucleotide sequence encoding a polypeptide having the specific properties as defined herein operably linked to a promoter.

ORGANISM

[0282] The term "organism" in relation to the present invention includes any organism that could comprise a nucleotide sequence according to the present invention or a nucleotide sequence encoding for a polypeptide having the specific properties as defined herein and/or products obtained therefrom.

[0283] The term "transgenic organism" in relation to the present invention includes any organism that comprises a nucleotide sequence coding for a polypeptide having the specific properties as defined herein and/or the products obtained therefrom, and/or wherein a promoter can allow expression of the nucleotide sequence coding for a polypeptide having the specific properties as defined herein within the organism. Preferably the nucleotide sequence is incorporated in the genome of the organism.

[0284] The term "transgenic organism" does not cover native nucleotide coding sequences in their natural environment when they are under the control of their native promoter which is also in its natural environment.

[0285] Therefore, the transgenic organism of the present disclosure includes an organism comprising any one of, or combinations of, a nucleotide sequence coding for a polypeptide having the specific properties as defined herein, constructs as defined herein, vectors as defined herein, plasmids as defined herein, cells as defined herein, or the products thereof. For example the transgenic organism can also comprise a nucleotide sequence coding for a polypeptide having the specific properties as defined herein under the control of a promoter not associated with a sequence encoding a lipid acyltransferase in nature.

TRANSFORMATION OF HOST CELLS/ORGANISM

[0286] Teachings on the transformation of prokaryotic hosts are well documented in the art, for example see Sambrook et al (Molecular Cloning: A Laboratory Manual, 2nd edition, 1989, Cold Spring Harbor Laboratory Press). If a prokaryotic host is used then the nucleotide sequence may need to be suitably modified before transformation - such as by removal of introns.

[0287] Various methods are known for the transformation of *Bacillus* species.

SECRETION

[0288] Often, it is desirable for the polypeptide to be secreted from the expression host into the culture medium from, where the enzyme may be more easily recovered. According to the present invention, the secretion leader sequence may be selected on the basis of the desired expression host. Hybrid signal sequences may also be used with the context of the present invention.

[0289] Typical examples of secretion leader sequences not associated with a nucleotide sequence encoding a lipid acyltransferase in nature are those originating from the fungal amyloglucosidase (AG) gene (*glaA* - both 18 and 24 amino acid versions e.g. from *Aspergillus*), the α -factor gene (yeasts e.g. *Saccharomyces*, *Kluyveromyces* and *Hansenula*) or the α -amylase gene (*Bacillus*).

DETECTION

[0290] A variety of protocols for detecting and measuring the expression of the amino acid sequence are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA) and fluorescent activated cell sorting (FACTS).

[0291] A wide variety of labels and conjugation techniques are known by those skilled in the art and can be used in various nucleic and amino acid assays.

[0292] A number of companies such as Pharmacia Biotech (Piscataway, NJ), Promega (Madison, WI), and US Biochemical Corp (Cleveland, OH) supply commercial kits and protocols for these procedures.

[0293] Suitable reporter molecules or labels include those radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents as well as substrates, cofactors, inhibitors, magnetic particles and the like. Patents teaching the use of such labels include US-A-3,817,837; US-A-3,850,752; US-A-3,939,350; US-A-3,996,345; US-A-4,277,437; US-A-4,275,149 and US-A-4,366,241.

[0294] Also, recombinant immunoglobulins may be produced as shown in US-A-4,816,567.

10 FUSION PROTEINS

[0295] In the method of the present disclosure the lipid acyltransferase may be produced as a fusion protein, for example to aid in extraction and purification thereof. Examples of fusion protein partners include glutathione-S-transferase (GST), 6xHis, GAL4 (DNA binding and/or transcriptional activation domains) and β -galactosidase. It may also be convenient to include a proteolytic cleavage site between the fusion protein partner and the protein sequence of interest to allow removal of fusion protein sequences. Preferably the fusion protein will not hinder the activity of the protein sequence.

[0296] Gene fusion expression systems in *E. coli* have been reviewed in Curr. Opin. Biotechnol. (1995) 6(5):501-6.

[0297] In another embodiment of the disclosure, the amino acid sequence of a polypeptide having the specific properties as defined herein may be ligated to a non-native sequence to encode a fusion protein. For example, for screening of peptide libraries for agents capable of affecting the substance activity, it may be useful to encode a chimeric substance expressing a non-native epitope that is recognised by a commercially available antibody.

[0298] The invention will now be described, by way of example only, with reference to the following Figures and Examples.

Figure 1 shows the amino acid sequence of a mutant *Aeromonas salmonicida* mature lipid acyltransferase (GCAT) with a mutation of Asn80Asp (notably, amino acid 80 is in the mature sequence) (SEQ ID 16);

Figure 2 shows an amino acid sequence (SEQ ID No. 1) a lipid acyl transferase from *Aeromonas hydrophila* (ATCC #7965);

Figure 3 shows a pfam00657 consensus sequence from database version 6 (SEQ ID No. 2);

5 Figure 4 shows an amino acid sequence (SEQ ID No. 3) obtained from the organism *Aeromonas hydrophila* (P10480; GI:121051);

Figure 5 shows an amino acid sequence (SEQ ID No. 4) obtained from the organism *Aeromonas salmonicida* (AAG098404; GI:9964017);

Figure 6 shows an amino acid sequence (SEQ ID No. 5) obtained from the organism *Streptomyces coelicolor* A3(2) (Genbank accession number NP_631558);

10 Figure 7 shows an amino acid sequence (SEQ ID No. 6) obtained from the organism *Streptomyces coelicolor* A3(2) (Genbank accession number: CAC42140);

Figure 8 shows an amino acid sequence (SEQ ID No. 7) obtained from the organism *Saccharomyces cerevisiae* (Genbank accession number P41734);

15 Figure 9 shows an amino acid sequence (SEQ ID No. 8) obtained from the organism *Ralstonia* (Genbank accession number: AL646052);

Figure 10 shows SEQ ID No. 9. Scoe1 NCBI protein accession code CAB39707.1 GI:4539178 conserved hypothetical protein [*Streptomyces coelicolor* A3(2)];

Figure 11 shows an amino acid shown as SEQ ID No. 10. Scoe2 NCBI protein accession code CAC01477.1 GI:9716139 conserved hypothetical protein [*Streptomyces coelicolor* A3(2)];

20 Figure 12 shows an amino acid sequence (SEQ ID No. 11) Scoe3 NCBI protein accession code CABS8833.1 GI:7635996 putative secreted protein. [*Streptomyces coelicolor* A3(2)];

Figure 13 shows an amino acid sequence (SEQ ID No. 12) Scoe4 NCBI protein accession code CAB89450.1 GI:7672261 putative secreted protein. [*Streptomyces coelicolor* A3(2)];

25 Figure 14 shows an amino acid sequence (SEQ ID No. 13) Scoe5 NCBI protein accession code CAB62724.1 GI:6562793 putative lipoprotein [*Streptomyces coelicolor* A3(2)];

Figure 15 shows an amino acid sequence (SEQ ID No. 14) Srim1 NCBI protein accession code AAK84028.1 GI:15082088 GDSL-lipase [*Streptomyces rimosus*];

Figure 16 shows an amino acid sequence (SEQ ID No. 15) of a lipid acyltransferase from *Aeromonas salmonicida* subsp. *Salmonicida* (ATCC#14174);

- 5 Figure 17 shows SEQ ID No. 19. Scoe1 NCBI protein accession code CAB39707.1 GI:4539178 conserved hypothetical protein [*Streptomyces coelicolor* A3(2)];

Figure 18 shows an amino acid sequence (SEQ ID No. 25) of the fusion construct used for mutagenesis of the *Aeromonas hydrophila* lipid acyltransferase gene. The underlined amino acids is a xylanase signal peptide;

- 10 Figure 19 shows a polypeptide sequence of a lipid acyltransferase enzyme from *Streptomyces* (SEQ ID No. 26);

Figure 20 shows a polypeptide sequence of a lipid acyltransferase enzyme from *Thermobifida* (SEQ ID No. 27);

- 15 Figure 21 shows a polypeptide sequence of a lipid acyltransferase enzyme from *Thermobifida* (SEQ ID No. 28);

Figure 22 shows a polypeptide of a lipid acyltransferase enzyme from *Corynebacterium efficiens* GDSx 300 amino acid (SEQ ID No. 29);

Figure 23 shows a polypeptide of a lipid acyltransferase enzyme from *Novosphingobium aromaticivorans* GDSx 284 amino acid (SEQ ID No. 30);

- 20 Figure 24 shows a polypeptide of a lipid acyltransferase enzyme from *Streptomyces coelicolor* GDSx 269 aa (SEQ ID No. 31);

Figure 25 shows a polypeptide of a lipid acyltransferase enzyme from *Streptomyces avermitilis* \ GDSx 269 amino acid (SEQ ID No. 32);

Figure 26 shows a polypeptide of a lipid acyltransferase enzyme from *Streptomyces* (SEQ ID No. 33);

- 25 Figure 27 shows an amino acid sequence (SEQ ID No. 34) obtained from the organism *Aeromonas hydrophila* (P10480; GI:121051) (notably, this is the mature sequence);

Figure 28 shows the amino acid sequence (SEQ ID No. 35) of a mutant *Aeromonas salmonicida* mature lipid acyltransferase (GCAT) (notably, this is the mature sequence);

Figure 29 shows a nucleotide sequence (SEQ ID No. 36) from *Streptomyces thermosacchari*;

Figure 30 shows an amino acid sequence (SEQ ID No. 37) from *Streptomyces thermosacchari*;

5 Figure 31 shows an amino acid sequence (SEQ ID No. 38) from *Thermobifida fusca*/GDSx 548 amino acid;

Figure 32 shows a nucleotide sequence (SEQ ID No. 39) from *Thermobifida fusca*;

Figure 33 shows an amino acid sequence (SEQ ID No. 40) from *Thermobifida fusca*/GDSx;

10 Figure 34 shows an amino acid sequence (SEQ ID No. 41) from *Corynebacterium efficiens*/GDSx 300 amino acid;

Figure 35 shows a nucleotide sequence (SEQ ID No. 42) from *Corynebacterium efficiens*;

Figure 36 shows an amino acid sequence (SEQ ID No. 43) from *S. coelicolor*/ GDSx 268 amino acid;

Figure 37 shows a nucleotide sequence (SEQ ID No. 44) from *S. coelicolor*;

Figure 38 shows an amino acid sequence (SEQ ID No. 45) from *S. avermitilis*;

15 Figure 39 shows a nucleotide sequence (SEQ ID No. 46) from *S. avermitilis*;

Figure 40 shows an amino acid sequence (SEQ ID No. 47) from *Thermobifida fusca*/GDSx;

Figure 41 shows a nucleotide sequence (SEQ ID No. 48) from *Thermobifida fusca*/GDSx;

20 Figure 42 shows an alignment of the L131 and homologues from *S. avermitilis* and *T. fusca* illustrates that the conservation of the GDSx motif (GDSY in L131 and *S. avermitilis* and *T. fusca*), the GANDY box, which is either GGND A or GGNDL, and the HPT block (considered to be the conserved catalytic histidine). These three conserved blocks are highlighted;

Figure 43 shows SEQ ID No 17 which is the amino acid sequence of a lipid acyltransferase from *Candida parapsilosis*;

Figure 44 shows SEQ ID No 18 which is the amino acid sequence of a lipid acyltransferase from *Candida parapsilosis*;

Figure 45 shows a ribbon representation of the 1IVN.PDB crystal structure which has glycerol in the active site. The Figure was made using the Deep View Swiss-PDB viewer;

5 Figure 46 shows 1IVN.PDB Crystal Structure - Side View using Deep View Swiss-PDB viewer, with glycerol in active site - residues within 10Å of active site glycerol are coloured black;

Figure 47 shows 1IVN.PDB Crystal Structure - Top View using Deep View Swiss-PDB viewer, with glycerol in active site - residues within 10Å of active site glycerol are coloured black;

Figure 48 shows alignment 1;

10 Figure 49 shows alignment 2;

Figures 50 and 51 show an alignment of 1IVN to P10480 (P10480 is the database sequence for *A. hydrophila* enzyme), this alignment was obtained from the PFAM database and used in the model building process; and

15 Figure 52 shows an alignment where P10480 is the database sequence for *Aeromonas hydrophila*. This sequence is used for the model construction and the site selection. Note that the full protein (SEQ ID No. 25) is depicted, the mature protein (equivalent to SEQ ID No. 34) starts at residue 19. *A. sal* is *Aeromonas salmonicida* (SEQ ID No. 4) GDSX lipase, *A. hyd* is *Aeromonas hydrophila* (SEQ ID No. 34) GDSX lipase. The consensus sequence contains a * at the position of a difference between the listed sequences.

20 Figure 53 shows a gene construct used in Example 1;

Figure 54 shows a codon optimised gene construct (no. 052907) used in Example 1; and

Figure 55 shows the sequence of the XhoI insert containing the LAT-KLM3' precursor gene, the -35 and -10 boxes are underlined;

25 Figure 56 shows BML780-KLM3'CAP50 (comprising SEQ ID No. 16 - upper colony) and BML780 (the empty host strain -- lower colony) after 48h growth at 37°C on 1% tributyrin agar;

Figure 57 shows a nucleotide sequence from *Aeromonas salmonicida* (SEQ ID No. 49) including the signal sequence (preLAT - positions 1 to 87);

Figure 58 shows a nucleotide sequence (SEQ ID No. 50) encoding a lipid acyl transferase according to the present invention obtained from the organism *Aeromonas hydrophila*;

5 Figure 59 shows a nucleotide sequence (SEQ ID No. 51) encoding a lipid acyl transferase according to the present invention obtained from the organism *Aeromonas salmonicida*;

Figure 60 shows a nucleotide sequence (SEQ ID No. 52) encoding a lipid acyl transferase according to the present invention obtained from the organism *Streptomyces coelicolor* A3(2) (Genbank accession number NC_003888.1:8327480..8328367);

10 Figure 61 shows a nucleotide sequence (SEQ ID No. 53) encoding a lipid acyl transferase according to the present invention obtained from the organism *Streptomyces coelicolor* A3(2) (Genbank accession number AL939131.1:265480..266367);

15 Figure 62 shows a nucleotide sequence (SEQ ID No. 54) encoding a lipid acyl transferase according to the present invention obtained from the organism *Saccharomyces cerevisiae* (Genbank accession number Z75034);

Figure 63 shows a nucleotide sequence (SEQ ID No. 55) encoding a lipid acyl transferase according to the present invention obtained from the organism *Ralstonia*;

Figure 64 shows a nucleotide sequence shown as SEQ ID No. 56 encoding NCBI protein accession code CAB39707.1 GI:4539178 conserved hypothetical protein [*Streptomyces coelicolor* A3(2)];

20 Figure 65 shows a nucleotide sequence shown as SEQ ID No. 57 encoding Scoe2 NCBI protein accession code CAC01477.1 GI:9716139 conserved hypothetical protein [*Streptomyces coelicolor*A3(2)];

Figure 66 shows a nucleotide sequence shown as SEQ ID No. 58 encoding Scoe3 NCBI protein accession code CAB88833.1 GI:7635996 putative secreted protein. [*Streptomyces coelicolor* A3(2)];

25 Figure 67 shows a nucleotide sequence shown as SEQ ID No. 59 encoding Scoe4 NCBI protein accession code CAB89450.1 GI:7672261 putative secreted protein. [*Streptomyces coelicolor* A3(2)];

Figure 68 shows a nucleotide sequence shown as SEQ ID No. 60, encoding Scoe5 NCBI protein accession code CAB62724.1 GI:6562793 putative lipoprotein [*Streptomyces coelicolor* A3(2)];

Figure 69 shows a nucleotide sequence shown as SEQ ID No. 61 encoding Srim1 NCBI protein accession code AAK84028.1 GI:15082088 GDSL-lipase [*Streptomyces rimosus*];

5 Figure 70 shows a nucleotide sequence (SEQ ID No. 62) encoding a lipid acyltransferase from *Aeromonas hydrophila* (ATCC #7965);

Figure 71 shows a nucleotide sequence (SEQ ID No 63) encoding a lipid acyltransferase from *Aeromonas salmonicida* subsp. *Salmonicida* (ATCC#14174); and

10 Figure 72 shows a nucleotide sequence (SEQ ID No. 24) encoding an enzyme from *Aeromonas hydrophila* including a xylanase signal peptide.

EXAMPLE 1

Expression of KLM3' in *Bacillus licheniformis*

[0299] A nucleotide sequence (SEQ ID No. 49) encoding a lipid acyltransferases (SEQ. ID No. 16, hereinafter KLM3') was expressed in *Bacillus licheniformis* as a fusion protein with the signal peptide of *B. licheniformis* [alpha]-amylase (LAT) (see FIGS. 53 and 54). For optimal expression in *Bacillus*, a codon optimized gene construct (no. 052907) was ordered at Geneart (Geneart AG, Regensburg, Germany).

[0300] Construct no. 052907 contains an incomplete LAT promoter (only the -10 sequence) in front of the LAT-KLM3' precursor gene and the LAT transcription (Tlat) downstream of the LAT-KLM3' precursor gene (see FIGS 53 and 55). To create a *Xho*I fragment that contains the LAT-KLM3' precursor gene flanked by the complete LAT promoter at the 5' end and the LAT terminator at the 3' end, a PCR (polymerase chain reaction) amplification was performed with the primers Plat5XhoI_FW and EBS2XhoI_RV and gene construct 052907 as template.

25 Plat5XhoI_FW:

ccccgcctcgaggctttcttttgaagaaaatatagggaatgggtactgttaaaaattc
ggaatatattatacaatatcatatgtttcacattgaaagggg

EBS2XhoI_RV: tggaaatctcaaggtttatcctttacctgtctcc

[0301] PCR was performed on a thermocycler with Phusion High Fidelity DNA polymerase (Finnzymes OY, Espoo, Finland) according to the instructions of the manufacturer (annealing temperature of 55[deg.] C.).

[0302] The resulting PCR fragment was digested with restriction enzyme *Xho*I and ligated with T4 DNA
5 ligase into *Xho*I digested pICatH according to the instructions of the supplier (Invitrogen, Carlsbad, Calif. USA).

[0303] The ligation mixture was transformed into *B. subtilis* strain SC6.1 as described in U.S. Patent Application US20020182734 (International Publication WO 02/14490). The sequence of the *Xho*I insert
10 containing the LAT-KLM3' precursor gene was confirmed by DNA sequencing (BaseClear, Leiden, The Netherlands) and one of the correct plasmid clones was designated pICatH-KLM3'(ori1) (Figure 53). pICatH-KLM3'(ori1) was transformed into *B. licheniformis* strain BML780 (a derivative of BRA7 and BML612, see WO2005111203) at the permissive temperature (37[deg.] C.).

[0304] One neomycin resistant (neoR) and chloramphenicol resistant (CmR) transformant was selected and designated BML780(pICatH-KLM3'(ori1)). The plasmid in BML780(pICatH-KLM3'(ori1)) was integrated
15 into the catH region on the *B. licheniformis* genome by growing the strain at a non-permissive temperature (50[deg.] C) in medium with 5 [mu]g/ml chloramphenicol. One CmR resistant clone was selected and designated BML780-pICatH-KLM3'(ori1). BML780-pICatH-KLM3'(ori1) was grown again at the permissive
20 temperature for several generations without antibiotics to loop-out vector sequences and then one neomycin sensitive (neoS), CmR clone was selected. In this clone, vector sequences of pICatH on the chromosome are excised (including the neomycin resistance gene) and only the catH - LATKLM3' cassette is left. Next, the catH - LATKLM3' cassette on the chromosome was amplified by growing the strain in/on media with increasing concentrations of chloramphenicol. After various rounds of amplification, one clone (resistant
25 against 50 [mu]g/ml chloramphenicol) was selected and designated BML780-KLM3'CAP50. To verify KLM3'expression, BML780-KLM3'CAP50 and BML780 (the empty host strain) were grown for 48h at 37 [deg.] C on a Heart Infusion (Bacto) agar plate with 1% tributyrin. A clearing zone, indicative for lipid acyltransferase activity, was clearly visible around the colony of BML780-KLM3'CAP50 but not around the host strain BML780 (see Figure 56). This result shows that a substantial amount of KLM3' is expressed in *B.*
30 *licheniformis* strain BML780-KLM3'CAP50 and that these KLM3' molecules are functional.

COMPARATIVE EXAMPLE 1

Vector construct

[0305] The plasmid construct is pCS32new N80D, which is a pCCmini derivative carrying the sequence encoding the mature form of the native *Aeromonas salmonicida* Glycerophospholipid-cholesterol acyltransferase with a Asn to Asp substitution at position 80 (KLM3'), under control of the p32 promoter and with a CGTase signal sequence.

5

[0306] The host strain used for the expression, is in the *bacillus subtilis* OS21ΔAprE strain

[0307] The expression level is measured as transferase activity, expressed as % cholesterol esterified, calculated from the difference in free cholesterol in the reference sample and free cholesterol in the enzyme sample in reactions with PC (T_{PC}) as donor and cholesterol as acceptor molecule.

10

Culture conditions

[0308] 5 ml of LB broth (Casein enzymatic digest, 10 g/l; low-sodium Yeast extract, 5 g/l; Sodium Chloride, 5 g/l; Inert tableting aids, 2 g/l) supplemented with 50 mg/l kanamycin, was inoculated with a single colony and incubated at 30 °C for 6 hours at 205 rpm. 0.7 ml of this culture was used to inoculate 50 ml of SAS media (K_2HPO_4 , 10 g/l; MOPS (3-morpholinopropane sulfonic acid), 40 g/l; Sodium Chloride, 5 g/l; Antifoam (Sin 260), 5 drops/l; Soy flour degreased, 20 g/l; Biospringer 106 (100 % dw YE), 20 g/l) supplemented with 50 mg/l kanamycin and a solution of high maltose starch hydrolysates (60 g/l). Incubation was continued for 40 hours at 30 °C and 180 rpm before the culture supernatant was separated by centrifugation at 19000 rpm for 30 min. The supernatant was transferred into a clean tube and directly used for transferase activity measurement.

15

20

Preparation of substrates and enzymatic reaction

[0309] PC (Avanti Polar Lipids #441601) and cholesterol (Sigma C8503) was scaled in the ratio 9:1, dissolved in chloroform, and evaporated to dryness.

The substrate was prepared by dispersion of 3% PC:Cholesterol 9:1 in 50 mM Hepes buffer pH 7.

0.250 ml substrate solution was transferred into a 3 ml glass tube with screw lid. 0.025 ml culture supernatant was added and the mixture was incubated at 40 °C for 2 hours. A reference sample with water instead of enzyme was also prepared. Heating the reaction mixture in a boiling water bath for 10 minutes stopped the enzyme reaction. 2 ml of 99% ethanol was added to the reaction mixture before submitted to cholesterol assay analysis.

25

Cholesterol assay

30

[0310] 100 µl substrate containing 1.4 U/ml Cholesterol oxidase(SERVA Electrophoresis GmbH cat. No 17109), 0.4 mg/ml ABTS (Sigma A-1888), 6 U/ml Peroxidase (Sigma 6782) in 0.1 M Tris-HCl, pH 6.6 and 0.5 % Triton X-100 (Sigma X-100) was incubated at 37°C for 5 minutes before 5 µl enzyme reaction sample was added and mixed. The reaction mixture was incubated for further 5 minutes and OD₄₀₅ was measured.

- 5 The content of cholesterol was calculated from the analyses of standard solutions of cholesterol containing 0.4 mg/ml, 0.3 mg/ml, 0.20 mg/ml, 0.1 mg/ml, 0.05 mg/ml, and 0 mg/ml cholesterol in 99 % EtOH.

Results

[0311] The table shows the average of 8 separate expression cultures

Strain	T _{PC} ^a
OS21ΔAprE[pCS32new]	74.2±10.1 ^b
^a T _{PC} is the transferase activity, expressed as % cholesterol esterified, calculated from the difference in free cholesterol in the reference sample and free cholesterol in the enzyme sample in reactions with PC as donor molecule and cholesterol as acceptor molecule.	
^b Average of 8 separate expression cultures	

- 10 [0312] The invention is further described in the following numbered paragraphs;

1. 1. A method for the production of a lipid acyltransferase comprising the steps of:
 - 20 (i) providing a *Bacillus licheniformis* cell;
 - (ii) transforming the *Bacillus licheniformis* cell with a heterologous nucleotide sequence encoding a lipid acyltransferase; and
 - 15 (iii) expressing the lipid acyltransferase in the cell under the control of a promoter sequence.
2. 2. A method according to paragraph 1 wherein the promoter sequence is not natively associated with the nucleotide sequence encoding a lipid acyltransferase.
3. 3. A method according to paragraph 2, wherein a nucleotide sequence encoding a signal peptide is operably linked to said heterologous nucleotide sequence encoding a lipid acyltransferase.
- 20 4. 4. A method according to any one of the preceding paragraphs, wherein said method comprises the additional step of isolating/recovering the lipid acyltransferase.
5. 5. A method according to any one of the preceding paragraphs wherein said promoter sequence is homologous to the host cell.
- 25 6. 6. A method according to any one of the preceding paragraphs wherein said promoter sequence is selected from the group consisting of an α-amylase promoter sequence, a protease promoter sequence,

a subtilisin promoter sequence, a glutamic acid-specific protease promoter sequence and a levansucrase promoter sequence.

7. 7. A method according to any one of the preceding paragraphs wherein said promoter sequence is the α -amylase promoter sequence.

8. 8. A method according to any one of the preceding paragraphs wherein the nucleotide sequence encoding a lipid acyl transferase encodes a lipid acyltransferase that comprises a GDSx motif and/ or a GANDY motif.

9. 9. A method according to any one of the preceding paragraphs wherein the nucleotide sequence encoding a lipid acyltransferase encodes a lipid acyltransferase that is characterised as an enzyme which possesses acyl transferase activity and which comprises the amino acid sequence motif GDSX, wherein X is one or more of the following amino acid residues L, A, V, I, F, Y, H, Q, T, N, M or S.

10. 10. A method according to any one of the preceding paragraphs wherein the nucleotide sequence encoding a lipid acyltransferase encodes a lipid acyltransferase that is obtainable from an organism from one or more of the following genera: *Aeromonas*, *Streptomyces*, *Saccharomyces*, *Lactococcus*, *Mycobacterium*, *Streptococcus*, *Lactobacillus*, *Desulfitobacterium*, *Bacillus*, *Campylobacter*, *Vibrionaceae*, *Xylella*, *Sulfolobus*, *Aspergillus*, *Schizosaccharomyces*, *Listeria*, *Neisseria*, *Mesorhizobium*, *Ralstonia*, *Xanthomonas* and *Candida*.

11. 11. A method according to any one of the preceding paragraphs wherein the nucleotide sequence encoding a lipid acyltransferase encodes a lipid acyltransferase that is obtainable from an organism from the genus *Aeromonas*.

12. 12. A method according to any one of the preceding paragraphs the nucleotide sequence encoding a lipid acyltransferase encodes a lipid acyltransferase that comprises an aspartic acid residue at a position corresponding to N-80 in the amino acid sequence of the *Aeromonas hydrophila* lipid acyltransferase shown as SEQ ID No. 35.

13. 13. A method according to any one of the preceding paragraphs, wherein the nucleotide sequence encoding a lipid acyltransferase encodes a lipid acyltransferase that comprises the amino acid sequence shown as SEQ ID No. 16, or an amino acid sequence which has at least 75% homology thereto.

14. 14. A method according to any one of the preceding paragraphs wherein the nucleotide sequence encoding a lipid acyltransferase encodes a lipid acyltransferase that comprises the amino acid sequence shown as SEQ ID No. 16.

15. 15. A *Bacillus licheniformis* host cell comprising a nucleotide sequence encoding a heterologous lipid acyltransferase.

16. 16. A host cell according to paragraph 14 wherein the nucleotide sequence encoding a lipid acyltransferase encodes a lipid acyltransferase that comprises a GDSx motif and/ or a GANDY motif.

17. 17. A host cell according to paragraph 15 or paragraph 16 wherein the nucleotide sequence encoding a lipid acyltransferase encodes a lipid acyltransferase that is characterised as an enzyme which possesses acyltransferase activity and which comprises the amino acid sequence motif GDSX, wherein X is one or more of the following amino acid residues L, A, V, I, F, Y, H, Q, T, N, M or S.
- 5 18. 18. A host cell according to any one of paragraphs 15 to 17 wherein the lipid acyltransferase is obtainable from an organism from one or more of the following genera: *Aeromonas*, *Streptomyces*, *Saccharomyces*, *Lactococcus*, *Mycobacterium*, *Streptococcus*, *Lactobacillus*, *Desulfotobacterium*, *Bacillus*, *Campylobacter*, *Vibrionaceae*, *Xylella*, *Sulfolobus*, *Aspergillus*, *Schizosaccharomyces*, *Listeria*, *Neisseria*, *Mesorhizobium*, *Ralstonia*, *Xanthomonas* and *Candida*.
- 10 19. 19. A host cell according to any one of paragraphs 15 to 17 wherein the nucleotide sequence encoding a lipid acyltransferase encodes a lipid acyltransferase that is obtainable from an organism from the genus *Aeromonas*.
20. 20. A host cell according to any one of paragraphs 15 to 19 wherein the nucleotide sequence encoding a lipid acyltransferase encodes a lipid acyltransferase that comprises an aspartic acid residue at a position corresponding to N-80 in the amino acid sequence of the *Aeromonas hydrophila* lipid acyltransferase shown as SEQ ID No. 35.
- 15 21. 21. A host cell according to any one of paragraphs 15 to 20 wherein the nucleotide sequence encoding a lipid acyltransferase encodes a lipid acyltransferase that comprises the amino acid sequence shown as SEQ ID No. 16, or an amino acid sequence which has 75% or more homology thereto.
- 20 22. 22. A host cell according to any one of paragraphs 15 to 21 wherein the nucleotide sequence encoding a lipid acyltransferase encodes a lipid acyltransferase that comprises the amino acid sequence shown as SEQ ID No. 16.
23. 23. Use of a *Bacillus licheniformis* host cell in the production of a heterologous lipid acyltransferase.
24. 24. Use according to paragraph 23 wherein there is an increased expression as compared with expression in a *B. subtilis* host cell.
- 25 25. 25. An expression vector comprising a nucleotide sequence encoding a lipid acyltransferase operably linked to a promoter sequence homologous to *B. licheniformis*.
26. 26. An expression vector according to paragraph 25, wherein the promoter is not natively associated with the nucleotide sequence encoding a lipid acyltransferase.
- 30 27. 27. An expression vector according to paragraph 25 or paragraph 26 wherein a nucleotide sequence encoding a signal peptide is operably linked to said heterologous nucleotide sequence encoding a lipid acyltransferase.
28. 28. An expression vector according to any one of the preceding paragraphs wherein said promoter sequence is selected from the group consisting of an α -amylase promoter sequence, a protease

promoter sequence, a subtilisin promoter sequence, a glutamic acid-specific protease promoter sequence and a levansucrase promoter sequence.

29. 29. An expression vector according to any one of the preceding paragraphs wherein said promoter sequence is the α -amylase promoter sequence.

5 30. 30. A method for the production of a lipid acyltransferase comprising the steps of:

1. (i) providing a *Bacillus* cell, wherein the *Bacillus* cell is one other than *B. subtilis*;
2. (ii) transforming the *Bacillus* cell, wherein the *Bacillus* cell is one other than *B. subtilis* with a heterologous nucleotide sequence encoding a lipid acyltransferase; and
3. (iii) expressing the lipid acyltransferase in the cell under the control of a promoter sequence.

10 31. 31. A *Bacillus* host cell wherein the *Bacillus* cell is one other than *B. subtilis* comprising a nucleotide sequence encoding a heterologous lipid acyltransferase.

32. 32. Use of a *Bacillus* host cell wherein the *Bacillus* host cell is one other than *B. subtilis* in the production of a heterologous lipid acyltransferase.

33. 33. A use as substantially hereinbefore described with reference to the Examples and Figures.

15 34. 34. An expression vector as substantially hereinbefore described with reference to the Examples and Figures.

35. 35. A *Bacillus licheniformis* host cell as substantially hereinbefore described with reference to the Examples and Figures.

20 SEQUENCE LISTING

[0313]

<110> DANISCO A/S

<120> PROCESS

<130> P025001WO

25 <140> PCT/IB2007/000558

< 141> 2007-01-25

<160> 118

<170> PatentIn version 3.5

67

<210> 1

<211> 335

<212> PRT

<213> *Acromonas hydrophila*

5

<400> 1

Met Lys Lys Trp Phe Val Cys Leu Leu Gly Leu Val Ala Leu Thr Val
 1 5 10 15

Gln Ala Ala Asp Ser Arg Pro Ala Phe Ser Arg Ile Val Met Phe Gly
 20 25 30

Asp Ser Leu Ser Asp Thr Gly Lys Met Tyr Ser Lys Met Arg Gly Tyr
 35 40 45

Leu Pro Ser Ser Pro Pro Tyr Tyr Glu Gly Arg Phe Ser Asn Gly Pro
 50 55 60

Val Trp Leu Glu Gln Leu Thr Lys Gln Phe Pro Gly Leu Thr Ile Ala
 65 70 75 80

Asn Glu Ala Glu Gly Gly Ala Thr Ala Val Ala Tyr Asn Lys Ile Ser
 85 90 95

Trp Asn Pro Lys Tyr Gln Val Ile Asn Asn Leu Asp Tyr Glu Val Thr
 100 105 110

Gln Phe Leu Gln Lys Asp Ser Phe Lys Pro Asp Asp Leu Val Ile Leu
 115 120 125

Trp Val Gly Ala Asn Asp Tyr Leu Ala Tyr Gly Trp Asn Thr Glu Gln
 130 135 140

Asp Ala Lys Arg Val Arg Asp Ala Ile Ser Asp Ala Ala Asn Arg Met
 145 150 155 160

Val Leu Asn Gly Ala Lys Gln Ile Leu Leu Phe Asn Leu Pro Asp Leu
 165 170 175

68

Gly Gln Asn Pro Ser Ala Arg Ser Gln Lys Val Val Glu Ala Val Ser
 180 185 190

His Val Ser Ala Tyr His Asn Gln Leu Leu Leu Asn Leu Ala Arg Gln
 195 200 205

Leu Ala Pro Thr Gly Met Val Lys Leu Phe Glu Ile Asp Lys Gln Phe
 210 215 220

Ala Glu Met Leu Arg Asp Pro Gln Asn Phe Gly Leu Ser Asp Val Glu
 225 230 235 240

Asn Pro Cys Tyr Asp Gly Gly Tyr Val Trp Lys Pro Phe Ala Thr Arg
 245 250 255

Ser Val Ser Thr Asp Arg Gln Leu Ser Ala Phe Ser Pro Gln Glu Arg
 260 265 270

Leu Ala Ile Ala Gly Asn Pro Leu Leu Ala Gln Ala Val Ala Ser Pro
 275 280 285

Met Ala Arg Arg Ser Ala Ser Pro Leu Asn Cys Glu Gly Lys Met Phe
 290 295 300

Trp Asp Gln Val His Pro Thr Thr Val Val His Ala Ala Leu Ser Glu
 305 310 315 320

Arg Ala Ala Thr Phe Ile Ala Asn Gln Tyr Glu Phe Leu Ala His
 325 330 335

<210> 2

< 211> 361

< 212> PRT

< 213> Artificial

<220>

< 223> Consensus sequence

69

<400> 2

Ile Val Ala Phe Gly Asp Ser Leu Thr Asp Gly Glu Ala Tyr Tyr Gly
1 5 10 15

Asp Ser Asp Gly Gly Gly Trp Gly Ala Gly Leu Ala Asp Arg Leu Thr
20 25 30

Ala Leu Leu Arg Leu Arg Ala Arg Pro Arg Gly Val Asp Val Phe Asn
35 40 45

Arg Gly Ile Ser Gly Arg Thr Ser Asp Gly Arg Leu Ile Val Asp Ala
50 55 60

70

Leu Val Ala Leu Leu Phe Leu Ala Gln Ser Leu Gly Leu Pro Asn Leu
 65 70 75 80

Pro Pro Tyr Leu Ser Gly Asp Phe Leu Arg Gly Ala Asn Phe Ala Ser
 85 90 95

Ala Gly Ala Thr Ile Leu Pro Thr Ser Gly Pro Phe Leu Ile Gln Val
 100 105 110

Gln Phe Lys Asp Phe Lys Ser Gln Val Leu Glu Leu Arg Gln Ala Leu
 115 120 125

Gly Leu Leu Gln Glu Leu Leu Arg Leu Leu Pro Val Leu Asp Ala Lys
 130 135 140

Ser Pro Asp Leu Val Thr Ile Met Ile Gly Thr Asn Asp Leu Ile Thr
 145 150 155 160

Ser Ala Phe Phe Gly Pro Lys Ser Thr Glu Ser Asp Arg Asn Val Ser
 165 170 175

Val Pro Glu Phe Lys Asp Asn Leu Arg Gln Leu Ile Lys Arg Leu Arg
 180 185 190

Ser Asn Asn Gly Ala Arg Ile Ile Val Leu Ile Thr Leu Val Ile Leu
 195 200 205

Asn Leu Gly Pro Leu Gly Cys Leu Pro Leu Lys Leu Ala Leu Ala Leu
 210 215 220

Ala Ser Ser Lys Asn Val Asp Ala Ser Gly Cys Leu Glu Arg Leu Asn
 225 230 235 240

Glu Ala Val Ala Asp Phe Asn Glu Ala Leu Arg Glu Leu Ala Ile Ser
 245 250 255

Lys Leu Glu Asp Gln Leu Arg Lys Asp Gly Leu Pro Asp Val Lys Gly
 260 265 270

Ala Asp Val Pro Tyr Val Asp Leu Tyr Ser Ile Phe Gln Asp Leu Asp
 275 280 285

Gly Ile Gln Asn Pro Ser Ala Tyr Val Tyr Gly Phe Glu Thr Thr Lys
 290 295 300

Ala Cys Cys Gly Tyr Gly Gly Arg Tyr Asn Tyr Asn Arg Val Cys Gly
 305 310 315 320

71

Asn Ala Gly Leu Cys Asn Val Thr Ala Lys Ala Cys Asn Pro Ser Ser
325 330 335

Tyr Leu Leu Ser Phe Leu Phe Trp Asp Gly Phe His Pro Ser Glu Lys
340 345 350

Gly Tyr Lys Ala Val Ala Glu Ala Leu
355 360

<210> 3

<211> 335

<212> PRT

<213> *Aeromonas hydrophila*

5

<400> 3

Met Lys Lys Trp Phe Val Cys Leu Leu Gly Leu Val Ala Leu Thr Val
1 5 10 15

Gln Ala Ala Asp Ser Arg Pro Ala Phe Ser Arg Ile Val Met Phe Gly
20 25 30

Asp Ser Leu Ser Asp Thr Gly Lys Met Tyr Ser Lys Met Arg Gly Tyr
35 40 45

Leu Pro Ser Ser Pro Pro Tyr Tyr Glu Gly Arg Phe Ser Asn Gly Pro
50 55 60

Val Trp Leu Glu Gln Leu Thr Asn Glu Phe Pro Gly Leu Thr Ile Ala
65 70 75 80

Asn Glu Ala Glu Gly Gly Pro Thr Ala Val Ala Tyr Asn Lys Ile Ser
85 90 95

Trp Asn Pro Lys Tyr Gln Val Ile Asn Asn Leu Asp Tyr Glu Val Thr
100 105 110

Gln Phe Leu Gln Lys Asp Ser Phe Lys Pro Asp Asp Leu Val Ile Leu
115 120 125

Trp Val Gly Ala Asn Asp Tyr Leu Ala Tyr Gly Trp Asn Thr Glu Gln
130 135 140

Asp Ala Lys Arg Val Arg Asp Ala Ile Ser Asp Ala Ala Asn Arg Met
145 150 155 160

Val Leu Asn Gly Ala Lys Glu Ile Leu Leu Phe Asn Leu Pro Asp Leu
165 170 175

Gly Gln Asn Pro Ser Ala Arg Ser Gln Lys Val Val Glu Ala Ala Ser

180	185	190
His Val Ser Ala Tyr His Asn Gln Leu Leu Leu Asn Leu Ala Arg Gln		
195	200	205
Leu Ala Pro Thr Gly Met Val Lys Leu Phe Glu Ile Asp Lys Gln Phe		
210	215	220
Ala Glu Met Leu Arg Asp Pro Gln Asn Phe Gly Leu Ser Asp Gln Arg		
225	230	235
Asn Ala Cys Tyr Gly Gly Ser Tyr Val Trp Lys Pro Phe Ala Ser Arg		
245	250	255
Ser Ala Ser Thr Asp Ser Gln Leu Ser Ala Phe Asn Pro Gln Glu Arg		
260	265	270
Leu Ala Ile Ala Gly Asn Pro Leu Leu Ala Gln Ala Val Ala Ser Pro		
275	280	285
Met Ala Ala Arg Ser Ala Ser Thr Leu Asn Cys Glu Gly Lys Met Phe		
290	295	300
Trp Asp Gln Val His Pro Thr Thr Val Val His Ala Ala Leu Ser Glu		
305	310	315
Pro Ala Ala Thr Phe Ile Glu Ser Gln Tyr Glu Phe Leu Ala His		
325	330	335

<210> 4

<211> 336

<212> PRT

<213> Aeromonas salmonicida

<400> 4

Met Lys Lys Trp Phe Val Cys Leu Leu Gly Leu Ile Ala Leu Thr Val
1 5 10 15
Gln Ala Ala Asp Thr Arg Pro Ala Phe Ser Arg Ile Val Met Phe Gly
20 25 30
Asp Ser Leu Ser Asp Thr Gly Lys Met Tyr Ser Lys Met Arg Gly Tyr
35 40 45
Leu Pro Ser Ser Pro Pro Tyr Tyr Glu Gly Arg Phe Ser Asn Gly Pro
50 55 60
Val Trp Leu Glu Gln Leu Thr Lys Gln Phe Pro Gly Leu Thr Ile Ala
65 70 75 80

Asn Glu Ala Glu Gly Gly Ala Thr Ala Val Ala Tyr Asn Lys Ile Ser
85 90 95

Trp Asn Pro Lys Tyr Gln Val Tyr Asn Asn Leu Asp Tyr Glu Val Thr
100 105 110

Gln Phe Leu Gln Lys Asp Ser Phe Lys Pro Asp Asp Leu Val Ile Leu
115 120 125

Trp Val Gly Ala Asn Asp Tyr Leu Ala Tyr Gly Trp Asn Thr Glu Gln
130 135 140

Asp Ala Lys Arg Val Arg Asp Ala Ile Ser Asp Ala Ala Asn Arg Met
145 150 155 160

Val Leu Asn Gly Ala Lys Gln Ile Leu Leu Phe Asn Leu Pro Asp Leu
165 170 175

Gly Gln Asn Pro Ser Ala Arg Ser Gln Lys Val Val Glu Ala Val Ser
180 185 190

His Val Ser Ala Tyr His Asn Lys Leu Leu Leu Asn Leu Ala Arg Gln
195 200 205

Leu Ala Pro Thr Gly Met Val Lys Leu Phe Glu Ile Asp Lys Gln Phe
210 215 220

Ala Glu Met Leu Arg Asp Pro Gln Asn Phe Gly Leu Ser Asp Val Glu
225 230 235 240

Asn Pro Cys Tyr Asp Gly Gly Tyr Val Trp Lys Pro Phe Ala Thr Arg
245 250 255

Ser Val Ser Thr Asp Arg Gln Leu Ser Ala Phe Ser Pro Gln Glu Arg
260 265 270

Leu Ala Ile Ala Gly Asn Pro Leu Leu Ala Gln Ala Val Ala Ser Pro
275 280 285

Met Ala Arg Arg Ser Ala Ser Pro Leu Asn Cys Glu Gly Lys Met Phe
290 295 300

Trp Asp Gln Val His Pro Thr Thr Val Val His Ala Ala Leu Ser Glu
305 310 315 320

Arg Ala Ala Thr Phe Ile Glu Thr Gln Tyr Glu Phe Leu Ala His Gly
325 330 335

<210> 5

<211> 295

74

<212> PRT

<213> Streptomyces coelicolor

<400> 5

Met Pro Lys Pro Ala Leu Arg Arg Val Met Thr Ala Thr Val Ala Ala
1 5 10 15

Val Gly Thr Leu Ala Leu Gly Leu Thr Asp Ala Thr Ala His Ala Ala
20 25 30

Pro Ala Gln Ala Thr Pro Thr Leu Asp Tyr Val Ala Leu Gly Asp Ser
35 40 45

Tyr Ser Ala Gly Ser Gly Val Leu Pro Val Asp Pro Ala Asn Leu Leu
50 55 60

Cys Leu Arg Ser Thr Ala Asn Tyr Pro His Val Ile Ala Asp Thr Thr
65 70 75 80

Gly Ala Arg Leu Thr Asp Val Thr Cys Gly Ala Ala Gln Thr Ala Asp
85 90 95

Phe Thr Arg Ala Gln Tyr Pro Gly Val Ala Pro Gln Leu Asp Ala Leu
100 105 110

Gly Thr Gly Thr Asp Leu Val Thr Leu Thr Ile Gly Gly Asn Asp Asn
115 120 125

Ser Thr Phe Ile Asn Ala Ile Thr Ala Cys Gly Thr Ala Gly Val Leu
130 135 140

Ser Gly Gly Lys Gly Ser Pro Cys Lys Asp Arg His Gly Thr Ser Phe
145 150 155 160

Asp Asp Glu Ile Glu Ala Asn Thr Tyr Pro Ala Leu Lys Glu Ala Leu
165 170 175

Leu Gly Val Arg Ala Arg Ala Pro His Ala Arg Val Ala Ala Leu Gly
180 185 190

Tyr Pro Trp Ile Thr Pro Ala Thr Ala Asp Pro Ser Cys Phe Leu Lys
195 200 205

Leu Pro Leu Ala Ala Gly Asp Val Pro Tyr Leu Arg Ala Ile Gln Ala
210 215 220

His Leu Asn Asp Ala Val Arg Arg Ala Ala Glu Glu Thr Gly Ala Thr
225 230 235 240

75

Tyr Val Asp Phe Ser Gly Val Ser Asp Gly His Asp Ala Cys Glu Ala
245 250 255

Pro Gly Thr Arg Trp Ile Glu Pro Leu Leu Phe Gly His Ser Leu Val
260 265 270

Pro Val His Pro Asn Ala Leu Gly Glu Arg Arg Met Ala Glu His Thr
275 280 285

Met Asp Val Leu Gly Leu Asp
290 295

<210> 6

< 211> 295

5

< 212> PRT

< 213> Streptomyces coelicolor

<400> 6

Met Pro Lys Pro Ala Leu Arg Arg Val Met Thr Ala Thr Val Ala Ala
1 5 10 15

Val Gly Thr Leu Ala Leu Gly Leu Thr Asp Ala Thr Ala His Ala Ala
20 25 30

Pro Ala Gln Ala Thr Pro Thr Leu Asp Tyr Val Ala Leu Gly Asp Ser
35 40 45

Tyr Ser Ala Gly Ser Gly Val Leu Pro Val Asp Pro Ala Asn Leu Leu
50 55 60

Cys Leu Arg Ser Thr Ala Asn Tyr Pro His Val Ile Ala Asp Thr Thr
65 70 75 80

Gly Ala Arg Leu Thr Asp Val Thr Cys Gly Ala Ala Gln Thr Ala Asp
85 90 95

Phe Thr Arg Ala Gln Tyr Pro Gly Val Ala Pro Gln Leu Asp Ala Leu
100 105 110

Gly Thr Gly Thr Asp Leu Val Thr Leu Thr Ile Gly Gly Asn Asp Asn
115 120 125

Ser Thr Phe Ile Asn Ala Ile Thr Ala Cys Gly Thr Ala Gly Val Leu
130 135 140

Ser Gly Gly Lys Gly Ser Pro Cys Lys Asp Arg His Gly Thr Ser Phe
145 150 155 160

76

Asp Asp Glu Ile Glu Ala Asn Thr Tyr Pro Ala Leu Lys Glu Ala Leu
 165 170 175

Leu Gly Val Arg Ala Arg Ala Pro His Ala Arg Val Ala Ala Leu Gly
 180 185 190

Tyr Pro Trp Ile Thr Pro Ala Thr Ala Asp Pro Ser Cys Phe Leu Lys
 195 200 205

Leu Pro Leu Ala Ala Gly Asp Val Pro Tyr Leu Arg Ala Ile Gln Ala
 210 215 220

His Leu Asn Asp Ala Val Arg Arg Ala Ala Glu Glu Thr Gly Ala Thr
 225 230 235 240

Tyr Val Asp Phe Ser Gly Val Ser Asp Gly His Asp Ala Cys Glu Ala
 245 250 255

Pro Gly Thr Arg Trp Ile Glu Pro Leu Leu Phe Gly His Ser Leu Val
 260 265 270

Pro Val His Pro Asn Ala Leu Gly Glu Arg Arg Met Ala Glu His Thr
 275 280 285

Met Asp Val Leu Gly Leu Asp
 290 295

<210> 7

< 211> 238

< 212> PRT

< 213> *Saccharomyces cerevisiae*

77

<400> 7

Met Asp Tyr Glu Lys Phe Leu Leu Phe Gly Asp Ser Ile Thr Glu Phe
 1 5 10 15

Ala Phe Asn Thr Arg Pro Ile Glu Asp Gly Lys Asp Gln Tyr Ala Leu
 20 25 30

Gly Ala Ala Leu Val Asn Glu Tyr Thr Arg Lys Met Asp Ile Leu Gln
 35 40 45

Arg Gly Phe Lys Gly Tyr Thr Ser Arg Trp Ala Leu Lys Ile Leu Pro
 50 55 60

Glu Ile Leu Lys His Glu Ser Asn Ile Val Met Ala Thr Ile Phe Leu
 65 70 75 80

Gly Ala Asn Asp Ala Cys Ser Ala Gly Pro Gln Ser Val Pro Leu Pro
 85 90 95

Glu Phe Ile Asp Asn Ile Arg Gln Met Val Ser Leu Met Lys Ser Tyr
 100 105 110

His Ile Arg Pro Ile Ile Ile Gly Pro Gly Leu Val Asp Arg Glu Lys
 115 120 125

Trp Glu Lys Glu Lys Ser Glu Glu Ile Ala Leu Gly Tyr Phe Arg Thr
 130 135 140

Asn Glu Asn Phe Ala Ile Tyr Ser Asp Ala Leu Ala Lys Leu Ala Asn
 145 150 155 160

Glu Glu Lys Val Pro Phe Val Ala Leu Asn Lys Ala Phe Gln Gln Glu
 165 170 175

Gly Gly Asp Ala Trp Gln Gln Leu Leu Thr Asp Gly Leu His Phe Ser
 180 185 190

Gly Lys Gly Tyr Lys Ile Phe His Asp Glu Leu Leu Lys Val Ile Glu
 195 200 205

Thr Phe Tyr Pro Gln Tyr His Pro Lys Asn Met Gln Tyr Lys Leu Lys
 210 215 220

Asp Trp Arg Asp Val Leu Asp Asp Gly Ser Asn Ile Met Ser
 225 230 235

5

<210> 8

<211> 347

<212> PRT

<213> Ralstonia sp.

78

<400> 8

Met Asn Leu Arg Gln Trp Met Gly Ala Ala Thr Ala Ala Leu Ala Leu
1 5 10 15

Gly Leu Ala Ala Cys Gly Gly Gly Gly Thr Asp Gln Ser Gly Asn Pro
20 25 30

Asn Val Ala Lys Val Gln Arg Met Val Val Phe Gly Asp Ser Leu Ser
35 40 45

Asp Ile Gly Thr Tyr Thr Pro Val Ala Gln Ala Val Gly Gly Gly Lys
50 55 60

Phe Thr Thr Asn Pro Gly Pro Ile Trp Ala Glu Thr Val Ala Ala Gln
65 70 75 80

79

Leu Gly Val Thr Leu Thr Pro Ala Val Met Gly Tyr Ala Thr Ser Val
85 90 95

Gln Asn Cys Pro Lys Ala Gly Cys Phe Asp Tyr Ala Gln Gly Gly Ser
100 105 110

Arg Val Thr Asp Pro Asn Gly Ile Gly His Asn Gly Gly Ala Gly Ala
115 120 125

Leu Thr Tyr Pro Val Gln Gln Gln Leu Ala Asn Phe Tyr Ala Ala Ser
130 135 140

Asn Asn Thr Phe Asn Gly Asn Asn Asp Val Val Phe Val Leu Ala Gly
145 150 155 160

Ser Asn Asp Ile Phe Phe Trp Thr Thr Ala Ala Ala Thr Ser Gly Ser
165 170 175

Gly Val Thr Pro Ala Ile Ala Thr Ala Gln Val Gln Gln Ala Ala Thr
180 185 190

Asp Leu Val Gly Tyr Val Lys Asp Met Ile Ala Lys Gly Ala Thr Gln
195 200 205

Val Tyr Val Phe Asn Leu Pro Asp Ser Ser Leu Thr Pro Asp Gly Val
210 215 220

Ala Ser Gly Thr Thr Gly Gln Ala Leu Leu His Ala Leu Val Gly Thr
225 230 235 240

Phe Asn Thr Thr Leu Gln Ser Gly Leu Ala Gly Thr Ser Ala Arg Ile
245 250 255

Ile Asp Phe Asn Ala Gln Leu Thr Ala Ala Ile Gln Asn Gly Ala Ser
260 265 270

Phe Gly Phe Ala Asn Thr Ser Ala Arg Ala Cys Asp Ala Thr Lys Ile
275 280 285

Asn Ala Leu Val Pro Ser Ala Gly Gly Ser Ser Leu Phe Cys Ser Ala
290 295 300

Asn Thr Leu Val Ala Ser Gly Ala Asp Gln Ser Tyr Leu Phe Ala Asp
305 310 315 320

Gly Val His Pro Thr Thr Ala Gly His Arg Leu Ile Ala Ser Asn Val
325 330 335

Leu Ala Arg Leu Leu Ala Asp Asn Val Ala His

340

345

<210> 9

<211> 261

80

< 212> PRT

< 213> Streptomyces coelicolor

<400> 9

Met Ile Gly Ser Tyr Val Ala Val Gly Asp Ser Phe Thr Glu Gly Val
 1 5 10 15

Gly Asp Pro Gly Pro Asp Gly Ala Phe Val Gly Trp Ala Asp Arg Leu
 20 25 30

Ala Val Leu Leu Ala Asp Arg Arg Pro Glu Gly Asp Phe Thr Tyr Thr
 35 40 45

Asn Leu Ala Val Arg Gly Arg Leu Leu Asp Gln Ile Val Ala Glu Gln
 50 55 60

Val Pro Arg Val Val Gly Leu Ala Pro Asp Leu Val Ser Phe Ala Ala
 65 70 75 80

Gly Gly Asn Asp Ile Ile Arg Pro Gly Thr Asp Pro Asp Glu Val Ala
 85 90 95

Glu Arg Phe Glu Leu Ala Val Ala Ala Leu Thr Ala Ala Ala Gly Thr
 100 105 110

Val Leu Val Thr Thr Gly Phe Asp Thr Arg Gly Val Pro Val Leu Lys
 115 120 125

His Leu Arg Gly Lys Ile Ala Thr Tyr Asn Gly His Val Arg Ala Ile
 130 135 140

Ala Asp Arg Tyr Gly Cys Pro Val Leu Asp Leu Trp Ser Leu Arg Ser
 145 150 155 160

Val Gln Asp Arg Arg Ala Trp Asp Ala Asp Arg Leu His Leu Ser Pro
 165 170 175

Glu Gly His Thr Arg Val Ala Leu Arg Ala Gly Gln Ala Leu Gly Leu
 180 185 190

Arg Val Pro Ala Asp Pro Asp Gln Pro Trp Pro Pro Leu Pro Pro Arg
 195 200 205

Gly Thr Leu Asp Val Arg Arg Asp Asp Val His Trp Ala Arg Glu Tyr
 210 215 220

Leu Val Pro Trp Ile Gly Arg Arg Leu Arg Gly Glu Ser Ser Gly Asp
225 230 235 240

His Val Thr Ala Lys Gly Thr Leu Ser Pro Asp Ala Ile Lys Thr Arg
245 250 255

Ile Ala Ala Val Ala
260

- <210> 10
- < 211> 260
- < 212> PRT
- < 213> Streptomyces coelicolor

<400> 10

Met Gln Thr Asn Pro Ala Tyr Thr Ser Leu Val Ala Val Gly Asp Ser
 1 5 10 15

Phe Thr Glu Gly Met Ser Asp Leu Leu Pro Asp Gly Ser Tyr Arg Gly
 20 25 30

Trp Ala Asp Leu Leu Ala Thr Arg Met Ala Ala Arg Ser Pro Gly Phe
 35 40 45

Arg Tyr Ala Asn Leu Ala Val Arg Gly Lys Leu Ile Gly Gln Ile Val
 50 55 60

Asp Glu Gln Val Asp Val Ala Ala Ala Met Gly Ala Asp Val Ile Thr
 65 70 75 80

Leu Val Gly Gly Leu Asn Asp Thr Leu Arg Pro Lys Cys Asp Met Ala
 85 90 95

Arg Val Arg Asp Leu Leu Thr Gln Ala Val Glu Arg Leu Ala Pro His
 100 105 110

Cys Glu Gln Leu Val Leu Met Arg Ser Pro Gly Arg Gln Gly Pro Val
 115 120 125

Leu Glu Arg Phe Arg Pro Arg Met Glu Ala Leu Phe Ala Val Ile Asp
 130 135 140

Asp Leu Ala Gly Arg His Gly Ala Val Val Val Asp Leu Tyr Gly Ala
 145 150 155 160

Gln Ser Leu Ala Asp Pro Arg Met Trp Asp Val Asp Arg Leu His Leu
 165 170 175

Thr Ala Glu Gly His Arg Arg Val Ala Glu Ala Val Trp Gln Ser Leu

[illegible]

<210> 11

< 211 > 454

5 < 212> PRT

< 212> PRT

< 213> *Streptomyces coelicolor*

84

<400> 11

Met Thr Arg Gly Arg Asp Gly Gly Ala Gly Ala Pro Pro Thr Lys His
 1 5 10 15

Arg Ala Leu Leu Ala Ala Ile Val Thr Leu Ile Val Ala Ile Ser Ala
 20 25 30

Ala Ile Tyr Ala Gly Ala Ser Ala Asp Asp Gly Ser Arg Asp His Ala
 35 40 45

Leu Gln Ala Gly Gly Arg Leu Pro Arg Gly Asp Ala Ala Pro Ala Ser
 50 55 60

Thr Gly Ala Trp Val Gly Ala Trp Ala Thr Ala Pro Ala Ala Ala Glu
 65 70 75 80

Pro Gly Thr Glu Thr Thr Gly Leu Ala Gly Arg Ser Val Arg Asn Val
 85 90 95

Val His Thr Ser Val Gly Gly Thr Gly Ala Arg Ile Thr Leu Ser Asn
 100 105 110

Leu Tyr Gly Gln Ser Pro Leu Thr Val Thr His Ala Ser Ile Ala Leu
 115 120 125

Ala Ala Gly Pro Asp Thr Ala Ala Ala Ile Ala Asp Thr Met Arg Arg
 130 135 140

85

Leu Thr Phe Gly Gly Ser Ala Arg Val Ile Ile Pro Ala Gly Gly Gln
 145 150 155 160

Val Met Ser Asp Thr Ala Arg Leu Ala Ile Pro Tyr Gly Ala Asn Val
 165 170 175

Leu Val Thr Thr Tyr Ser Pro Ile Pro Ser Gly Pro Val Thr Tyr His
 180 185 190

Pro Gln Ala Arg Gln Thr Ser Tyr Leu Ala Asp Gly Asp Arg Thr Ala
 195 200 205

Asp Val Thr Ala Val Ala Tyr Thr Thr Pro Thr Pro Tyr Trp Arg Tyr
 210 215 220

Leu Thr Ala Leu Asp Val Leu Ser His Glu Ala Asp Gly Thr Val Val
 225 230 235 240

Ala Phe Gly Asp Ser Ile Thr Asp Gly Ala Arg Ser Gln Ser Asp Ala
 245 250 255

Asn His Arg Trp Thr Asp Val Leu Ala Ala Arg Leu His Glu Ala Ala
 260 265 270

Gly Asp Gly Arg Asp Thr Pro Arg Tyr Ser Val Val Asn Glu Gly Ile
 275 280 285

Ser Gly Asn Arg Leu Leu Thr Ser Arg Pro Gly Arg Pro Ala Asp Asn
 290 295 300

Pro Ser Gly Leu Ser Arg Phe Gln Arg Asp Val Leu Glu Arg Thr Asn
 305 310 315 320

Val Lys Ala Val Val Val Val Leu Gly Val Asn Asp Val Leu Asn Ser
 325 330 335

Pro Glu Leu Ala Asp Arg Asp Ala Ile Leu Thr Gly Leu Arg Thr Leu
 340 345 350

Val Asp Arg Ala His Ala Arg Gly Leu Arg Val Val Gly Ala Thr Ile
 355 360 365

Thr Pro Phe Gly Gly Tyr Gly Gly Tyr Thr Glu Ala Arg Glu Thr Met
 370 375 380

Arg Gln Glu Val Asn Glu Glu Ile Arg Ser Gly Arg Val Phe Asp Thr
 385 390 395 400

86

Val Val Asp Phe Asp Lys Ala Leu Arg Asp Pro Tyr Asp Pro Arg Arg
405 410 415

Met Arg Ser Asp Tyr Asp Ser Gly Asp His Leu His Pro Gly Asp Lys
420 425 430

Gly Tyr Ala Arg Met Gly Ala Val Ile Asp Leu Ala Ala Leu Lys Gly
435 440 445

Ala Ala Pro Val Lys Ala
450

<210> 12

<211> 340

<212> PRT

5 <213> Streptomyces coelicolor

<400> 12

Met Thr Ser Met Ser Arg Ala Arg Val Ala Arg Arg Ile Ala Ala Gly
1 5 10 15

Ala Ala Tyr Gly Gly Gly Gly Ile Gly Leu Ala Gly Ala Ala Val
20 25 30

Gly Leu Val Val Ala Glu Val Gln Leu Ala Arg Arg Arg Val Gly Val
35 40 45

Gly Thr Pro Thr Arg Val Pro Asn Ala Gln Gly Leu Tyr Gly Gly Thr
50 55 60

Leu Pro Thr Ala Gly Asp Pro Pro Leu Arg Leu Met Met Leu Gly Asp
65 70 75 80

Ser Thr Ala Ala Gly Gln Gly Val His Arg Ala Gly Gln Thr Pro Gly
85 90 95

Ala Leu Leu Ala Ser Gly Leu Ala Ala Val Ala Glu Arg Pro Val Arg
100 105 110

Leu Gly Ser Val Ala Gln Pro Gly Ala Cys Ser Asp Asp Leu Asp Arg
115 120 125

Gln Val Ala Leu Val Leu Ala Glu Pro Asp Arg Val Pro Asp Ile Cys
130 135 140

Val Ile Met Val Gly Ala Asn Asp Val Thr His Arg Met Pro Ala Thr
145 150 155 160

Arg Ser Val Arg His Leu Ser Ser Ala Val Arg Arg Leu Arg Thr Ala
165 170 175

87

Gly Ala Glu Val Val Val Gly Thr Cys Pro Asp Leu Gly Thr Ile Glu
180 185 190

Arg Val Arg Gln Pro Leu Arg Trp Leu Ala Arg Arg Ala Ser Arg Gln
195 200 205

Leu Ala Ala Ala Gln Thr Ile Gly Ala Val Glu Gln Gly Gly Arg Thr
210 215 220

Val Ser Leu Gly Asp Leu Leu Gly Pro Glu Phe Ala Gln Asn Pro Arg
225 230 235 240

Glu Leu Phe Gly Pro Asp Asn Tyr His Pro Ser Ala Glu Gly Tyr Ala
245 250 255

Thr Ala Ala Met Ala Val Leu Pro Ser Val Cys Ala Ala Leu Gly Leu
260 265 270

Trp Pro Ala Asp Glu Glu His Pro Asp Ala Leu Arg Arg Glu Gly Phe
275 280 285

Leu Pro Val Ala Arg Ala Ala Ala Glu Ala Ala Ser Glu Ala Gly Thr
290 295 300

Glu Val Ala Ala Ala Met Pro Thr Gly Pro Arg Gly Pro Trp Ala Leu
305 310 315 320

Leu Lys Arg Arg Arg Arg Arg Arg Val Ser Glu Ala Glu Pro Ser Ser
325 330 335

Pro Ser Gly Val
340

<210> 13

<211> 305

<212> PRT

<213> Streptomyces coelicolor

<400> 13

Met Gly Arg Gly Thr Asp Gln Arg Thr Arg Tyr Gly Arg Arg Arg Ala
1 5 10 15

Arg Val Ala Leu Ala Ala Leu Thr Ala Ala Val Leu Gly Val Gly Val
20 25 30

Ala Gly Cys Asp Ser Val Gly Gly Asp Ser Pro Ala Pro Ser Gly Ser
35 40 45

88

Pro Ser Lys Arg Thr Arg Thr Ala Pro Ala Trp Asp Thr Ser Pro Ala
50 55 60

Ser Val Ala Ala Val Gly Asp Ser Ile Thr Arg Gly Phe Asp Ala Cys
65 70 75 80

Ala Val Leu Ser Asp Cys Pro Glu Val Ser Trp Ala Thr Gly Ser Ser
85 90 95

Ala Lys Val Asp Ser Leu Ala Val Arg Leu Leu Gly Lys Ala Asp Ala
100 105 110

Ala Glu His Ser Trp Asn Tyr Ala Val Thr Gly Ala Arg Met Ala Asp
115 120 125

Leu Thr Ala Gln Val Thr Arg Ala Ala Gln Arg Glu Pro Glu Leu Val
130 135 140

Ala Val Met Ala Gly Ala Asn Asp Ala Cys Arg Ser Thr Thr Ser Ala
145 150 155 160

Met Thr Pro Val Ala Asp Phe Arg Ala Gln Phe Glu Glu Ala Met Ala
165 170 175

Thr Leu Arg Lys Lys Leu Pro Lys Ala Gln Val Tyr Val Ser Ser Ile
180 185 190

Pro Asp Leu Lys Arg Leu Trp Ser Gln Gly Arg Thr Asn Pro Leu Gly
195 200 205

Lys Gln Val Trp Lys Leu Gly Leu Cys Pro Ser Met Leu Gly Asp Ala
210 215 220

Asp Ser Leu Asp Ser Ala Ala Thr Leu Arg Arg Asn Thr Val Arg Asp
225 230 235 240

Arg Val Ala Asp Tyr Asn Glu Val Leu Arg Glu Val Cys Ala Lys Asp
245 250 255

Arg Arg Cys Arg Ser Asp Asp Gly Ala Val His Glu Phe Arg Phe Gly
260 265 270

Thr Asp Gln Leu Ser His Trp Asp Trp Phe His Pro Ser Val Asp Gly
275 280 285

Gln Ala Arg Leu Ala Glu Ile Ala Tyr Arg Ala Val Thr Ala Lys Asn
290 295 300

Pro

305

<210> 14

<211> 268

< 212> PRT

< 213> Streptomyces rimosus

<400> 14

Met Arg Leu Ser Arg Arg Ala Ala Thr Ala Ser Ala Leu Leu Leu Thr
 1 5 10 15

Pro Ala Leu Ala Leu Phe Gly Ala Ser Ala Ala Val Ser Ala Pro Arg
 20 25 30

Ile Gln Ala Thr Asp Tyr Val Ala Leu Gly Asp Ser Tyr Ser Ser Gly
 35 40 45

Val Gly Ala Gly Ser Tyr Asp Ser Ser Ser Gly Ser Cys Lys Arg Ser
 50 55 60

Thr Lys Ser Tyr Pro Ala Leu Trp Ala Ala Ser His Thr Gly Thr Arg
 65 70 75 80

Phe Asn Phe Thr Ala Cys Ser Gly Ala Arg Thr Gly Asp Val Leu Ala
 85 90 95

Lys Gln Leu Thr Pro Val Asn Ser Gly Thr Asp Leu Val Ser Ile Thr
 100 105 110

Ile Gly Gly Asn Asp Ala Gly Phe Ala Asp Thr Met Thr Thr Cys Asn
 115 120 125

Leu Gln Gly Glu Ser Ala Cys Leu Ala Arg Ile Ala Lys Ala Arg Ala
 130 135 140

Tyr Ile Gln Gln Thr Leu Pro Ala Gln Leu Asp Gln Val Tyr Asp Ala
 145 150 155 160

Ile Asp Ser Arg Ala Pro Ala Ala Gln Val Val Val Leu Gly Tyr Pro
 165 170 175

Arg Phe Tyr Lys Leu Gly Gly Ser Cys Ala Val Gly Leu Ser Glu Lys
 180 185 190

Ser Arg Ala Ala Ile Asn Ala Ala Ala Asp Asp Ile Asn Ala Val Thr
 195 200 205

Ala Lys Arg Ala Ala Asp His Gly Phe Ala Phe Gly Asp Val Asn Thr
 210 215 220

Thr Phe Ala Gly His Glu Leu Cys Ser Gly Ala Pro Trp Leu His Ser
225 230 235 240

Val Thr Leu Pro Val Glu Asn Ser Tyr His Pro Thr Ala Asn Gly Gln
245 250 255

Ser Lys Gly Tyr Leu Pro Val Leu Asn Ser Ala Thr
260 265

<210> 15

< 211> 336

< 212> PRT

< 213> Aeromonas salmonicida

<400> 15

Met Lys Lys Trp Phe Val Cys Leu Leu Gly Leu Ile Ala Leu Thr Val
 1 5 10 15

Gln Ala Ala Asp Thr Arg Pro Ala Phe Ser Arg Ile Val Met Phe Gly
 20 25 30

Asp Ser Leu Ser Asp Thr Gly Lys Met Tyr Ser Lys Met Arg Gly Tyr
 35 40 45

Leu Pro Ser Ser Pro Pro Tyr Tyr Glu Gly Arg Phe Ser Asn Gly Pro
 50 55 60

Val Trp Leu Glu Gln Leu Thr Lys Gln Phe Pro Gly Leu Thr Ile Ala
 65 70 75 80

Asn Glu Ala Glu Gly Gly Ala Thr Ala Val Ala Tyr Asn Lys Ile Ser
 85 90 95

Trp Asn Pro Lys Tyr Gln Val Ile Asn Asn Leu Asp Tyr Glu Val Thr
 100 105 110

Gln Phe Leu Gln Lys Asp Ser Phe Lys Pro Asp Asp Leu Val Ile Leu
 115 120 125

Trp Val Gly Ala Asn Asp Tyr Leu Ala Tyr Gly Trp Asn Thr Glu Gln
 130 135 140

Asp Ala Lys Arg Val Arg Asp Ala Ile Ser Asp Ala Ala Asn Arg Met
 145 150 155 160

Val Leu Asn Gly Ala Lys Gln Ile Leu Leu Phe Asn Leu Pro Asp Leu
 165 170 175

Gly Gln Asn Pro Ser Ala Arg Ser Gln Lys Val Val Glu Ala Val Ser

	180		185		190	
His Val Ser Ala Tyr His Asn Lys Leu Leu Leu Asn Leu Ala Arg Gln	195		200		205	
Leu Ala Pro Thr Gly Met Val Lys Leu Phe Glu Ile Asp Lys Gln Phe	210		215		220	
Ala Glu Met Leu Arg Asp Pro Gln Asn Phe Gly Leu Ser Asp Val Glu	225		230		235	240
Asn Pro Cys Tyr Asp Gly Gly Tyr Val Trp Lys Pro Phe Ala Thr Arg		245		250		255
Ser Val Ser Thr Asp Arg Gln Leu Ser Ala Phe Ser Pro Gln Glu Arg		260		265		270
Leu Ala Ile Ala Gly Asn Pro Leu Leu Ala Gln Ala Val Ala Ser Pro		275		280		285
Met Ala Arg Arg Ser Ala Ser Pro Leu Asn Cys Glu Gly Lys Met Phe		290		295		300
Trp Asp Gln Val His Pro Thr Thr Val Val His Ala Ala Leu Ser Glu		305		310		315
						320
Arg Ala Ala Thr Phe Ile Glu Thr Gln Tyr Glu Phe Leu Ala His Gly		325		330		335

<210> 16

<211> 318

<212> PRT

<213> Aeromonas salmonicida

<400> 16

Ala Asp Thr Arg Pro Ala Phe Ser Arg Ile Val Met Phe Gly Asp Ser
1 5 10 15

Leu Ser Asp Thr Gly Lys Met Tyr Ser Lys Met Arg Gly Tyr Leu Pro
20 25 30

Ser Ser Pro Pro Tyr Tyr Glu Gly Arg Phe Ser Asn Gly Pro Val Trp
35 40 45

Leu Glu Gln Leu Thr Lys Gln Phe Pro Gly Leu Thr Ile Ala Asn Glu
50 55 60

Ala Glu Gly Gly Ala Thr Ala Val Ala Tyr Asn Lys Ile Ser Trp Asp
65 70 75 80

94

Pro Lys Tyr Gln Val Ile Asn Asn Leu Asp Tyr Glu Val Thr Gln Phe
85 90 95

Leu Gln Lys Asp Ser Phe Lys Pro Asp Asp Leu Val Ile Leu Trp Val
100 105 110

Gly Ala Asn Asp Tyr Leu Ala Tyr Gly Trp Asn Thr Glu Gln Asp Ala
115 120 125

Lys Arg Val Arg Asp Ala Ile Ser Asp Ala Ala Asn Arg Met Val Leu
130 135 140

Asn Gly Ala Lys Gln Ile Leu Leu Phe Asn Leu Pro Asp Leu Gly Gln
145 150 155 160

Asn Pro Ser Ala Arg Ser Gln Lys Val Val Glu Ala Val Ser His Val
165 170 175

Ser Ala Tyr His Asn Lys Leu Leu Leu Asn Leu Ala Arg Gln Leu Ala
180 185 190

Pro Thr Gly Met Val Lys Leu Phe Glu Ile Asp Lys Gln Phe Ala Glu
195 200 205

Met Leu Arg Asp Pro Gln Asn Phe Gly Leu Ser Asp Val Glu Asn Pro
210 215 220

Cys Tyr Asp Gly Gly Tyr Val Trp Lys Pro Phe Ala Thr Arg Ser Val
225 230 235 240

Ser Thr Asp Arg Gln Leu Ser Ala Phe Ser Pro Gln Glu Arg Leu Ala
245 250 255

Ile Ala Gly Asn Pro Leu Leu Ala Gln Ala Val Ala Ser Pro Met Ala
260 265 270

Arg Arg Ser Ala Ser Pro Leu Asn Cys Glu Gly Lys Met Phe Trp Asp
275 280 285

Gln Val His Pro Thr Thr Val Val His Ala Ala Leu Ser Glu Arg Ala
290 295 300

Ala Thr Phe Ile Glu Thr Gln Tyr Glu Phe Leu Ala His Gly
305 310 315

<210> 17

<211> 465

<212> PRT

<213> Candida parapsilosis

<400> 17

Met Arg Tyr Phe Ala Ile Ala Phe Leu Leu Ile Asn Thr Ile Ser Ala
 1 5 10 15

Phe Val Leu Ala Pro Lys Lys Pro Ser Gln Asp Asp Phe Tyr Thr Pro
 20 25 30

Pro Gln Gly Tyr Glu Ala Gln Pro Leu Gly Ser Ile Leu Lys Thr Arg
 35 40 45

Asn Val Pro Asn Pro Leu Thr Asn Val Phe Thr Pro Val Lys Val Gln
 50 55 60

Asn Ala Trp Gln Leu Leu Val Arg Ser Glu Asp Thr Phe Gly Asn Pro
 65 70 75 80

Asn Ala Ile Val Thr Thr Ile Ile Gln Pro Phe Asn Ala Lys Lys Asp
 85 90 95

Lys Leu Val Ser Tyr Gln Thr Phe Glu Asp Ser Gly Lys Leu Asp Cys
 100 105 110

Ala Pro Ser Tyr Ala Ile Gln Tyr Gly Ser Asp Ile Ser Thr Leu Thr
 115 120 125

Thr Gln Gly Glu Met Tyr Tyr Ile Ser Ala Leu Leu Asp Gln Gly Tyr
 130 135 140

Tyr Val Val Thr Pro Asp Tyr Glu Gly Pro Lys Ser Thr Phe Thr Val
 145 150 155 160

Gly Leu Gln Ser Gly Arg Ala Thr Leu Asn Ser Leu Arg Ala Thr Leu
 165 170 175

Lys Ser Gly Asn Leu Thr Gly Val Ser Ser Asp Ala Glu Thr Leu Leu
 180 185 190

Trp Gly Tyr Ser Gly Gly Ser Leu Ala Ser Gly Trp Ala Ala Ala Ile
 195 200 205

Gln Lys Glu Tyr Ala Pro Glu Leu Ser Lys Asn Leu Leu Gly Ala Ala
 210 215 220

Leu Gly Gly Phe Val Thr Asn Ile Thr Ala Thr Ala Glu Ala Val Asp
 225 230 235 240

Ser Gly Pro Phe Ala Gly Ile Ile Ser Asn Ala Leu Ala Gly Ile Gly
 245 250 255

Asn Glu Tyr Pro Asp Phe Lys Asn Tyr Leu Leu Lys Lys Val Ser Pro
260 265 270

Leu Leu Ser Ile Thr Tyr Arg Leu Gly Asn Thr His Cys Leu Leu Asp
275 280 285

Gly Gly Ile Ala Tyr Phe Gly Lys Ser Phe Phe Ser Arg Ile Ile Arg
290 295 300

Tyr Phe Pro Asp Gly Trp Asp Leu Val Asn Gln Glu Pro Ile Lys Thr
305 310 315 320

Ile Leu Gln Asp Asn Gly Leu Val Tyr Gln Pro Lys Asp Leu Thr Pro
325 330 335

Gln Ile Pro Leu Phe Ile Tyr His Gly Thr Leu Asp Ala Ile Val Pro
340 345 350

Ile Val Asn Ser Arg Lys Thr Phe Gln Gln Trp Cys Asp Trp Gly Leu
355 360 365

Lys Ser Gly Glu Tyr Asn Glu Asp Leu Thr Asn Gly His Ile Thr Glu
370 375 380

Ser Ile Val Gly Ala Pro Ala Ala Leu Thr Trp Ile Ile Asn Arg Phe
385 390 395 400

Asn Gly Gln Pro Pro Val Asp Gly Cys Gln His Asn Val Arg Ala Ser
405 410 415

Asn Leu Glu Tyr Pro Gly Thr Pro Gln Ser Ile Lys Asn Tyr Phe Glu
420 425 430

Ala Ala Leu His Ala Ile Leu Gly Phe Asp Leu Gly Pro Asp Val Lys
435 440 445

Arg Asp Lys Val Thr Leu Gly Gly Leu Leu Lys Leu Glu Arg Phe Ala
450 455 460

Phe
465

<210> 18

< 211> 471

< 212> PRT

< 213> Candida parapsilosis

<400> 18

Met Arg Tyr Phe Ala Ile Ala Phe Leu Ieu Ile Asn Thr Ile Ser Ala
1 5 10 15

Phe Val Leu Ala Pro Lys Lys Pro Ser Gln Asp Asp Phe Tyr Thr Pro
20 25 30

Pro Gln Gly Tyr Glu Ala Gln Pro Leu Gly Ser Ile Leu Lys Thr Arg
35 40 45

Asn Val Pro Asn Pro Leu Thr Asn Val Phe Thr Pro Val Lys Val Gln
50 55 60

Asn Ala Trp Gln Leu Leu Val Arg Ser Glu Asp Thr Phe Gly Asn Pro
65 70 75 80

Asn Ala Ile Val Thr Thr Ile Ile Gln Pro Phe Asn Ala Lys Lys Asp
85 90 95

Lys Leu Val Ser Tyr Gln Thr Phe Glu Asp Ser Gly Lys Leu Asp Cys
100 105 110

Ala Pro Ser Tyr Ala Ile Gln Tyr Gly Ser Asp Ile Ser Thr Leu Thr
115 120 125

Thr Gln Gly Glu Met Tyr Tyr Ile Ser Ala Leu Leu Asp Gln Gly Tyr
130 135 140

Tyr Val Val Thr Pro Asp Tyr Glu Gly Pro Lys Ser Thr Phe Thr Val
145 150 155 160

Gly Leu Gln Ser Gly Arg Ala Thr Leu Asn Ser Leu Arg Ala Thr Leu
165 170 175

Lys Ser Gly Asn Leu Thr Gly Val Ser Ser Asp Ala Glu Thr Leu Leu
180 185 190

Trp Gly Tyr Ser Gly Gly Ser Leu Ala Ser Gly Trp Ala Ala Ala Ile
195 200 205

Gln Lys Glu Tyr Ala Pro Glu Leu Ser Lys Asn Leu Leu Gly Ala Ala
210 215 220

Leu Gly Gly Phe Val Thr Asn Ile Thr Ala Thr Ala Glu Ala Val Asp
225 230 235 240

Ser Gly Pro Phe Ala Gly Ile Ile Ser Asn Ala Leu Ala Gly Ile Gly
245 250 255

Asn Glu Tyr Pro Asp Phe Lys Asn Tyr Leu Leu Lys Lys Val Ser Pro

[illegible]

<210> 19

< 211> 261

< 212> PRT

< 213> *Streptomyces coelicolor*

<400> 19

Met Ile Gly Ser Tyr Val Ala Val Gly Asp Ser Phe Thr Glu Gly Val
1 5 10 15

Gly Asp Pro Gly Pro Asp Gly Ala Phe Val Gly Trp Ala Asp Arg Leu
 20 25 30
 Ala Val Leu Leu Ala Asp Arg Arg Pro Glu Gly Asp Phe Thr Tyr Thr
 35 40 45
 Asn Leu Ala Val Arg Gly Arg Leu Leu Asp Gln Ile Val Ala Glu Gln
 50 55 60
 Val Pro Arg Val Val Gly Leu Ala Pro Asp Leu Val Ser Phe Ala Ala
 65 70 75 80
 Gly Gly Asn Asp Ile Ile Arg Pro Gly Thr Asp Pro Asp Glu Val Ala
 85 90 95
 Glu Arg Phe Glu Leu Ala Val Ala Ala Leu Thr Ala Ala Ala Gly Thr
 100 105 110
 Val Leu Val Thr Thr Gly Phe Asp Thr Arg Gly Val Pro Val Leu Lys
 115 120 125
 His Leu Arg Gly Lys Ile Ala Thr Tyr Asn Gly His Val Arg Ala Ile
 130 135 140
 Ala Asp Arg Tyr Gly Cys Pro Val Leu Asp Leu Trp Ser Leu Arg Ser
 145 150 155 160
 Val Gln Asp Arg Arg Ala Trp Asp Ala Asp Arg Leu His Leu Ser Pro
 165 170 175
 Glu Gly His Thr Arg Val Ala Leu Arg Ala Gly Gln Ala Leu Gly Leu
 180 185 190
 Arg Val Pro Ala Asp Pro Asp Gln Pro Trp Pro Pro Leu Pro Pro Arg
 195 200 205
 Gly Thr Leu Asp Val Arg Arg Asp Asp Val His Trp Ala Arg Glu Tyr
 210 215 220
 Leu Val Pro Trp Ile Gly Arg Arg Leu Arg Gly Glu Ser Ser Gly Asp
 225 230 235 240
 His Val Thr Ala Lys Gly Thr Leu Ser Pro Asp Ala Ile Lys Thr Arg
 245 250 255
 Ile Ala Ala Val Ala
 260

<210> 20

<211> 230

100

< 212> PRT

< 213> *Aspergillus aculeatus*

<400> 20

Thr Thr Val Tyr Leu Ala Gly Asp Ser Thr Met Ala Lys Asn Gly Gly
 1 5 10 15

Gly Ser Gly Thr Asn Gly Trp Gly Glu Tyr Leu Ala Ser Tyr Leu Ser
 20 25 30

Ala Thr Val Val Asn Asp Ala Val Ala Gly Arg Ser Ala Arg Ser Tyr
 35 40 45

Thr Arg Glu Gly Arg Phe Glu Asn Ile Ala Asp Val Val Thr Ala Gly
 50 55 60

Asp Tyr Val Ile Val Glu Phe Gly His Asn Asp Gly Gly Ser Leu Ser
 65 70 75 80

Thr Asp Asn Gly Arg Thr Asp Cys Ser Gly Thr Gly Ala Glu Val Cys
 85 90 95

Tyr Ser Val Tyr Asp Gly Val Asn Glu Thr Ile Leu Thr Phe Pro Ala
 100 105 110

Tyr Leu Glu Asn Ala Ala Lys Leu Phe Thr Ala Lys Gly Ala Lys Val
 115 120 125

Ile Leu Ser Ser Gln Thr Pro Asn Asn Pro Trp Glu Thr Gly Thr Phe
 130 135 140

Val Asn Ser Pro Thr Arg Phe Val Glu Tyr Ala Glu Leu Ala Ala Glu
 145 150 155 160

Val Ala Gly Val Glu Tyr Val Asp His Trp Ser Tyr Val Asp Ser Ile
 165 170 175

Tyr Glu Thr Leu Gly Asn Ala Thr Val Asn Ser Tyr Phe Pro Ile Asp
 180 185 190

His Thr His Thr Ser Pro Ala Gly Ala Glu Val Val Ala Glu Ala Phe
 195 200 205

Leu Lys Ala Val Val Cys Thr Gly Thr Ser Leu Lys Ser Val Leu Thr
 210 215 220

Thr Thr Ser Phe Glu Gly
 225 230

101

<210> 21

< 211> 18

< 212> PRT

< 213> Aeromonas sp.

5

<400> 21

Met	Lys	Lys	Trp	Phe	Val	Cys	Leu	Leu	Gly	Leu	Ile	Ala	Leu	Thr	Val
1				5					10					15	

Gln Ala

<210> 22

< 211> 29

< 212> PRT

10

< 213> Bacillus subtilis

<400> 22

Met	Arg	Ser	Lys	Lys	Leu	Trp	Ile	Ser	Leu	Leu	Phe	Ala	Leu	Thr	Leu
1				5					10					15	

Ile	Phe	Thr	Met	Ala	Phe	Ser	Asn	Met	Ser	Ala	Gln	Ala
			20					25				

<210> 23

< 211> 29

15

< 212> PRT

< 213> Bacillus licheniformis

<400> 23

Met	Met	Arg	Lys	Lys	Ser	Phe	Trp	Phe	Gly	Met	Leu	Thr	Ala	Phe	Met
1				5					10					15	

Leu	Val	Phe	Thr	Met	Glu	Phe	Ser	Asp	Ser	Ala	Ser	Ala
			20					25				

<210> 24

20

< 211> 1047

< 212> DNA

< 213> Aeromonas hydrophila

<400> 24

atgtttaagt ttaaaaagaa tttcttagtt ggattatcgg cagctttaat gagtattagc	60
ttgttttcgg caaccgcctc tgcagctagc gccgacagcc gtcccgccctt ttcccgatc	120
gtgatgttcg gcgacagcct ctccgatacc ggcaaaatgt acageaagat gcgcgggttac	180
ctcccctcca gcccgcccta ctatgagggc cgtttctcca acggaccctt ctggctggag	240
cagctgacca aacagttccc gggctctgacc atcgccaacg aagcggaagg cggcgccact	300
gccgtggcctt acaacaagat ctcttgggaat cccaagtatc aggtcatcaa caacctggac	360
tacgaggtca cccagttctt gcagaaagac agcttcaagc cggacgatct ggtgacctc	420
tgggtcgggtg ccaatgacta tctggcctat ggctggaaca cggagcagga tgccaagcgg	480
gttcgcgatg ccatcagcga tgcggccaac cgcctgttac tgaacgggtc caagcagata	540
ctgctgttca acctgcgga tctgggccag aaccctcag ctgcagtc gaaggtggtc	600
gaggcgggtc gccatgtctc cgcctatcac aaccagctgc tgcgtgaacct ggcacgccag	660
ctggccccca ccggcatggt aaagctgttc gagatcgaca agcaatttgc cgagatgctg	720
cgtgatccgc agaacttcgg cctgagcgac gtcgagaacc cctgctacga cggcggctat	780
gtgtggaagc cgtttgccac ccgcagcgtc agcaccgacc gccagctctc cgccttcagt	840
ccgcaggaaac gcctcgccat ccgcggcaac ccgctgctgg cacaggccgt tgccagtcct	900
atggcccgcc gcagcgccag cccctcaac tgtgagggca agatgttctg ggatcaggta	960
cacccgacca ctgtcgtgca cgcagccctg agcgagcgcg ccgccacctt catcggaac	1020
cagtaacgagt tcctcgccca ctgatga	1047

5

<210> 25

<211> 347

<212> PRT

<213> Artificial

<220>

10

<223> Fusion construct

103

<400> 25

Met Phe Lys Phe Lys Lys Asn Phe Leu Val Gly Leu Ser Ala Ala Leu
 1 5 10 15

Met Ser Ile Ser Leu Phe Ser Ala Thr Ala Ser Ala Ala Ser Ala Asp
 20 25 30

Ser Arg Pro Ala Phe Ser Arg Ile Val Met Phe Gly Asp Ser Leu Ser
 35 40 45

Asp Thr Gly Lys Met Tyr Ser Lys Met Arg Gly Tyr Leu Pro Ser Ser
 50 55 60

Pro Pro Tyr Tyr Glu Gly Arg Phe Ser Asn Gly Pro Val Trp Leu Glu
 65 70 75 80

Gln Leu Thr Lys Gln Phe Pro Gly Leu Thr Ile Ala Asn Glu Ala Glu
 85 90 95

Gly Gly Ala Thr Ala Val Ala Tyr Asn Lys Ile Ser Trp Asn Pro Lys
 100 105 110

Tyr Gln Val Ile Asn Asn Leu Asp Tyr Glu Val Thr Gln Phe Leu Gln
 115 120 125

104

Lys Asp Ser Phe Lys Pro Asp Asp Leu Val Ile Leu Trp Val Gly Ala
130 135 140

Asn Asp Tyr Leu Ala Tyr Gly Trp Asn Thr Glu Gln Asp Ala Lys Arg
145 150 155 160

Val Arg Asp Ala Ile Ser Asp Ala Ala Asn Arg Met Val Leu Asn Gly
165 170 175

Ala Lys Gln Ile Leu Leu Phe Asn Leu Pro Asp Leu Gly Gln Asn Pro
180 185 190

Ser Ala Arg Ser Gln Lys Val Val Glu Ala Val Ser His Val Ser Ala
195 200 205

Tyr His Asn Gln Leu Leu Leu Asn Leu Ala Arg Gln Leu Ala Pro Thr
210 215 220

Gly Met Val Lys Leu Phe Glu Ile Asp Lys Gln Phe Ala Glu Met Leu
225 230 235 240

Arg Asp Pro Gln Asn Phe Gly Leu Ser Asp Val Glu Asn Pro Cys Tyr
245 250 255

Asp Gly Gly Tyr Val Trp Lys Pro Phe Ala Thr Arg Ser Val Ser Thr
260 265 270

Asp Arg Gln Leu Ser Ala Phe Ser Pro Gln Glu Arg Leu Ala Ile Ala
275 280 285

Gly Asn Pro Leu Leu Ala Gln Ala Val Ala Ser Pro Met Ala Arg Arg
290 295 300

Ser Ala Ser Pro Leu Asn Cys Glu Gly Lys Met Phe Trp Asp Gln Val
305 310 315 320

His Pro Thr Thr Val Val His Ala Ala Leu Ser Glu Arg Ala Ala Thr
325 330 335

Phe Ile Ala Asn Gln Tyr Glu Phe Leu Ala His
340 345

<210> 26

<211> 267

<212> PRT

<213> Streptomyces sp.

105

<400> 26

Met Arg Leu Thr Arg Ser Leu Ser Ala Ala Ser Val Ile Val Phe Ala
1 5 10 15

Leu Leu Leu Ala Leu Leu Gly Ile Ser Pro Ala Gln Ala Ala Gly Pro
20 25 30

Ala Tyr Val Ala Leu Gly Asp Ser Tyr Ser Ser Gly Asn Gly Ala Gly
35 40 45

Ser Tyr Ile Asp Ser Ser Gly Asp Cys His Arg Ser Asn Asn Ala Tyr
50 55 60

Pro Ala Arg Trp Ala Ala Ala Asn Ala Pro Ser Ser Phe Thr Phe Ala
65 70 75 80

Ala Cys Ser Gly Ala Val Thr Thr Asp Val Ile Asn Asn Gln Leu Gly
85 90 95

Ala Leu Asn Ala Ser Thr Gly Leu Val Ser Ile Thr Ile Gly Gly Asn
100 105 110

Asp Ala Gly Phe Ala Asp Ala Met Thr Thr Cys Val Thr Ser Ser Asp
115 120 125

Ser Thr Cys Leu Asn Arg Leu Ala Thr Ala Thr Asn Tyr Ile Asn Thr
130 135 140

Thr Leu Leu Ala Arg Leu Asp Ala Val Tyr Ser Gln Ile Lys Ala Arg
145 150 155 160

Ala Pro Asn Ala Arg Val Val Val Leu Gly Tyr Pro Arg Met Tyr Leu
165 170 175

Ala Ser Asn Pro Trp Tyr Cys Leu Gly Leu Ser Asn Thr Lys Arg Ala
180 185 190

Ala Ile Asn Thr Thr Ala Asp Thr Leu Asn Ser Val Ile Ser Ser Arg
195 200 205

Ala Thr Ala His Gly Phe Arg Phe Gly Asp Val Arg Pro Thr Phe Asn
210 215 220

Asn His Glu Leu Phe Phe Gly Asn Asp Trp Leu His Ser Leu Thr Leu
225 230 235 240

Pro Val Trp Glu Ser Tyr His Pro Thr Ser Thr Gly His Gln Ser Gly
245 250 255

Tyr Leu Pro Val Leu Asn Ala Asn Ser Ser Thr

260

265

106

<210> 27

<211> 548

<212> PRT

<213> Thermobifida sp.

5

<400> 27

Met Leu Pro His Pro Ala Gly Glu Arg Gly Glu Val Gly Ala Phe Phe
 1 5 10 15

Ala Leu Leu Val Gly Thr Pro Gln Asp Arg Arg Leu Arg Leu Glu Cys
 20 25 30

His Glu Thr Arg Pro Leu Arg Gly Arg Cys Gly Cys Gly Glu Arg Arg
 35 40 45

Val Pro Pro Leu Thr Leu Pro Gly Asp Gly Val Leu Cys Thr Thr Ser
 50 55 60

Ser Thr Arg Asp Ala Glu Thr Val Trp Arg Lys His Leu Gln Pro Arg
 65 70 75 80

Pro Asp Gly Gly Phe Arg Pro His Leu Gly Val Gly Cys Leu Leu Ala
 85 90 95

Gly Gln Gly Ser Pro Gly Val Leu Trp Cys Gly Arg Glu Gly Cys Arg
 100 105 110

Phe Glu Val Cys Arg Arg Asp Thr Pro Gly Leu Ser Arg Thr Arg Asn
 115 120 125

Gly Asp Ser Ser Pro Pro Phe Arg Ala Gly Trp Ser Leu Pro Pro Lys
 130 135 140

Cys Gly Glu Ile Ser Gln Ser Ala Arg Lys Thr Pro Ala Val Pro Arg
 145 150 155 160

Tyr Ser Leu Leu Arg Thr Asp Arg Pro Asp Gly Pro Arg Gly Arg Phe
 165 170 175

Val Gly Ser Gly Pro Arg Ala Ala Thr Arg Arg Arg Leu Phe Leu Gly
 180 185 190

Ile Pro Ala Leu Val Leu Val Thr Ala Leu Thr Leu Val Leu Ala Val
 195 200 205

Pro Thr Gly Arg Glu Thr Leu Trp Arg Met Trp Cys Glu Ala Thr Gln
 210 215 220

Asp Trp Cys Leu Gly Val Pro Val Asp Ser Arg Gly Gln Pro Ala Glu
 225 230 235 240
 Asp Gly Glu Phe Leu Leu Leu Ser Pro Val Gln Ala Ala Thr Trp Gly
 245 250 255
 Asn Tyr Tyr Ala Leu Gly Asp Ser Tyr Ser Ser Gly Asp Gly Ala Arg
 260 265 270
 Asp Tyr Tyr Pro Gly Thr Ala Val Lys Gly Gly Cys Trp Arg Ser Ala
 275 280 285
 Asn Ala Tyr Pro Glu Leu Val Ala Glu Ala Tyr Asp Phe Ala Gly His
 290 295 300
 Leu Ser Phe Leu Ala Cys Ser Gly Gln Arg Gly Tyr Ala Met Leu Asp
 305 310 315 320
 Ala Ile Asp Glu Val Gly Ser Gln Leu Asp Trp Asn Ser Pro His Thr
 325 330 335
 Ser Leu Val Thr Ile Gly Ile Gly Gly Asn Asp Leu Gly Phe Ser Thr
 340 345 350
 Val Leu Lys Thr Cys Met Val Arg Val Pro Leu Leu Asp Ser Lys Ala
 355 360 365
 Cys Thr Asp Gln Glu Asp Ala Ile Arg Lys Arg Met Ala Lys Phe Glu
 370 375 380
 Thr Thr Phe Glu Glu Leu Ile Ser Glu Val Arg Thr Arg Ala Pro Asp
 385 390 395 400
 Ala Arg Ile Leu Val Val Gly Tyr Pro Arg Ile Phe Pro Glu Glu Pro
 405 410 415
 Thr Gly Ala Tyr Tyr Thr Leu Thr Ala Ser Asn Gln Arg Trp Leu Asn
 420 425 430
 Glu Thr Ile Gln Glu Phe Asn Gln Gln Leu Ala Glu Ala Val Ala Val
 435 440 445
 His Asp Glu Glu Ile Ala Ala Ser Gly Gly Val Gly Ser Val Glu Phe
 450 455 460
 Val Asp Val Tyr His Ala Leu Asp Gly His Glu Ile Gly Ser Asp Glu
 465 470 475 480

108

Pro Trp Val Asn Gly Val Gln Leu Arg Asp Leu Ala Thr Gly Val Thr
485 490 495

Val Asp Arg Ser Thr Phe His Pro Asn Ala Ala Gly His Arg Ala Val
500 505 510

Gly Glu Arg Val Ile Glu Gln Ile Glu Thr Gly Pro Gly Arg Pro Leu
515 520 525

Tyr Ala Thr Phe Ala Val Val Ala Gly Ala Thr Val Asp Thr Leu Ala
530 535 540

Gly Glu Val Gly
545

<210> 28

< 211> 372

< 212> PRT

5 < 213> Thermobifida sp.

<400> 28

Met Gly Ser Gly Pro Arg Ala Ala Thr Arg Arg Arg Leu Phe Leu Gly
1 5 10 15

Ile Pro Ala Leu Val Leu Val Thr Ala Leu Thr Leu Val Leu Ala Val
20 25 30

Pro Thr Gly Arg Glu Thr Leu Trp Arg Met Trp Cys Glu Ala Thr Gln
35 40 45

Asp Trp Cys Leu Gly Val Pro Val Asp Ser Arg Gly Gln Pro Ala Glu
50 55 60

Asp Gly Glu Phe Leu Leu Leu Ser Pro Val Gln Ala Ala Thr Trp Gly
65 70 75 80

Asn Tyr Tyr Ala Leu Gly Asp Ser Tyr Ser Ser Gly Asp Gly Ala Arg
85 90 95

Asp Tyr Tyr Pro Gly Thr Ala Val Lys Gly Gly Cys Trp Arg Ser Ala
100 105 110

Asn Ala Tyr Pro Glu Leu Val Ala Glu Ala Tyr Asp Phe Ala Gly His
115 120 125

Leu Ser Phe Leu Ala Cys Ser Gly Gln Arg Gly Tyr Ala Met Leu Asp
130 135 140

Ala Ile Asp Glu Val Gly Ser Gln Leu Asp Trp Asn Ser Pro His Thr
145 150 155 160

109

Ser Leu Val Thr Ile Gly Ile Gly Gly Asn Asp Leu Gly Phe Ser Thr
165 170 175

Val Leu Lys Thr Cys Met Val Arg Val Pro Leu Leu Asp Ser Lys Ala
180 185 190

Cys Thr Asp Gln Glu Asp Ala Ile Arg Lys Arg Met Ala Lys Phe Glu
195 200 205

Thr Thr Phe Glu Glu Leu Ile Ser Glu Val Arg Thr Arg Ala Pro Asp
210 215 220

Ala Arg Ile Leu Val Val Gly Tyr Pro Arg Ile Phe Pro Glu Glu Pro
225 230 235 240

Thr Gly Ala Tyr Tyr Thr Leu Thr Ala Ser Asn Gln Arg Trp Leu Asn
245 250 255

Glu Thr Ile Gln Glu Phe Asn Gln Gln Leu Ala Glu Ala Val Ala Val
260 265 270

His Asp Glu Glu Ile Ala Ala Ser Gly Gly Val Gly Ser Val Glu Phe
275 280 285

Val Asp Val Tyr His Ala Leu Asp Gly His Glu Ile Gly Ser Asp Glu
290 295 300

Pro Trp Val Asn Gly Val Gln Leu Arg Asp Leu Ala Thr Gly Val Thr
305 310 315 320

Val Asp Arg Ser Thr Phe His Pro Asn Ala Ala Gly His Arg Ala Val
325 330 335

Gly Glu Arg Val Ile Glu Gln Ile Glu Thr Gly Pro Gly Arg Pro Leu
340 345 350

Tyr Ala Thr Phe Ala Val Val Ala Gly Ala Thr Val Asp Thr Leu Ala
355 360 365

Gly Glu Val Gly
370

<210> 29

<211> 300

<212> PRT

<213> Corynebacterium efficiens

<400> 29

```

Met Arg Thr Thr Val Ile Ala Ala Ser Ala Leu Leu Leu Leu Ala Gly
1          5          10          15

Cys Ala Asp Gly Ala Arg Glu Glu Thr Ala Gly Ala Pro Pro Gly Glu
20          25          30

Ser Ser Gly Gly Ile Arg Glu Glu Gly Ala Glu Ala Ser Thr Ser Ile
35          40          45

Thr Asp Val Tyr Ile Ala Leu Gly Asp Ser Tyr Ala Ala Met Gly Gly
50          55          60

Arg Asp Gln Pro Leu Arg Gly Glu Pro Phe Cys Leu Arg Ser Ser Gly
65          70          75          80

Asn Tyr Pro Glu Leu Leu His Ala Glu Val Thr Asp Leu Thr Cys Gln
85          90          95

Gly Ala Val Thr Gly Asp Leu Leu Glu Pro Arg Thr Leu Gly Glu Arg
100         105         110

Thr Leu Pro Ala Gln Val Asp Ala Leu Thr Glu Asp Thr Thr Leu Val
115         120         125

Thr Leu Ser Ile Gly Gly Asn Asp Leu Gly Phe Gly Glu Val Ala Gly
130         135         140

Cys Ile Arg Glu Arg Ile Ala Gly Glu Asn Ala Asp Asp Cys Val Asp
145         150         155         160

Leu Leu Gly Glu Thr Ile Gly Glu Gln Leu Asp Gln Leu Pro Pro Gln
165         170         175

Leu Asp Arg Val His Glu Ala Ile Arg Asp Arg Ala Gly Asp Ala Gln
180         185         190

Val Val Val Thr Gly Tyr Leu Pro Leu Val Ser Ala Gly Asp Cys Pro
195         200         205

Glu Leu Gly Asp Val Ser Glu Ala Asp Arg Arg Trp Ala Val Glu Leu
210         215         220

Thr Gly Gln Ile Asn Glu Thr Val Arg Glu Ala Ala Glu Arg His Asp
225         230         235         240

Ala Leu Phe Val Leu Pro Asp Asp Ala Asp Glu His Thr Ser Cys Ala
245         250         255

Pro Pro Gln Gln Arg Trp Ala Asp Ile Gln Gly Gln Gln Thr Asp Ala

```

111

260 265 270

Tyr Pro Leu His Pro Thr Ser Ala Gly His Glu Ala Met Ala Ala Ala
275 280 285

Val Arg Asp Ala Leu Gly Leu Glu Pro Val Gln Pro
290 295 300

- <210> 30
- < 211> 284
- < 212> PRT
- < 213> Novosphingobium aromaticivorans

<400> 30

Met Gly Gln Val Lys Leu Phe Ala Arg Arg Cys Ala Pro Val Leu Leu
 1 5 10 15

Ala Leu Ala Gly Leu Ala Pro Ala Ala Thr Val Ala Arg Glu Ala Pro
 20 25 30

Leu Ala Glu Gly Ala Arg Tyr Val Ala Leu Gly Ser Ser Phe Ala Ala
 35 40 45

Gly Pro Gly Val Gly Pro Asn Ala Pro Gly Ser Pro Glu Arg Cys Gly
 50 55 60

Arg Gly Thr Leu Asn Tyr Pro His Leu Leu Ala Glu Ala Leu Lys Leu
 65 70 75 80

Asp Leu Val Asp Ala Thr Cys Ser Gly Ala Thr Thr His His Val Leu
 85 90 95

Gly Pro Trp Asn Glu Val Pro Pro Gln Ile Asp Ser Val Asn Gly Asp
 100 105 110

Thr Arg Leu Val Thr Leu Thr Ile Gly Gly Asn Asp Val Ser Phe Val
 115 120 125

Gly Asn Ile Phe Ala Ala Ala Cys Glu Lys Met Ala Ser Pro Asp Pro
 130 135 140

Arg Cys Gly Lys Trp Arg Glu Ile Thr Glu Glu Glu Trp Gln Ala Asp
 145 150 155 160

Glu Glu Arg Met Arg Ser Ile Val Arg Gln Ile His Ala Arg Ala Pro
 165 170 175

Leu Ala Arg Val Val Val Val Asp Tyr Ile Thr Val Leu Pro Pro Ser
 180 185 190

113

Gly Thr Cys Ala Ala Met Ala Ile Ser Pro Asp Arg Leu Ala Gln Ser
195 200 205

Arg Ser Ala Ala Lys Arg Leu Ala Arg Ile Thr Ala Arg Val Ala Arg
210 215 220

Glu Glu Gly Ala Ser Leu Leu Lys Phe Ser His Ile Ser Arg Arg His
225 230 235 240

His Pro Cys Ser Ala Lys Pro Trp Ser Asn Gly Leu Ser Ala Pro Ala
245 250 255

Asp Asp Gly Ile Pro Val His Pro Asn Arg Leu Gly His Ala Glu Ala
260 265 270

Ala Ala Ala Leu Val Lys Leu Val Lys Leu Met Lys
275 280

<210> 31

<211> 268

5

<212> PRT

<213> Streptomyces coelicolor

<400> 31

Met Arg Arg Phe Arg Leu Val Gly Phe Leu Ser Ser Leu Val Leu Ala
1 5 10 15

Ala Gly Ala Ala Leu Thr Gly Ala Ala Thr Ala Gln Ala Ala Gln Pro
20 25 30

Ala Ala Ala Asp Gly Tyr Val Ala Leu Gly Asp Ser Tyr Ser Ser Gly
35 40 45

Val Gly Ala Gly Ser Tyr Ile Ser Ser Ser Gly Asp Cys Lys Arg Ser
50 55 60

Thr Lys Ala His Pro Tyr Leu Trp Ala Ala Ala His Ser Pro Ser Thr
65 70 75 80

Phe Asp Phe Thr Ala Cys Ser Gly Ala Arg Thr Gly Asp Val Leu Ser
85 90 95

Gly Gln Leu Gly Pro Leu Ser Ser Gly Thr Gly Leu Val Ser Ile Ser
100 105 110

Ile Gly Gly Asn Asp Ala Gly Phe Ala Asp Thr Met Thr Thr Cys Val
115 120 125

Leu Gln Ser Glu Ser Ser Cys Leu Ser Arg Ile Ala Thr Ala Glu Ala

114

130	135	140
Tyr Val Asp Ser Thr Leu Pro Gly Lys Leu Asp Gly Val Tyr Ser Ala		
145	150	155 160
Ile Ser Asp Lys Ala Pro Asn Ala His Val Val Val Ile Gly Tyr Pro		
	165	170 175
Arg Phe Tyr Lys Leu Gly Thr Thr Cys Ile Gly Leu Ser Glu Thr Lys		
	180	185 190
Arg Thr Ala Ile Asn Lys Ala Ser Asp His Leu Asn Thr Val Leu Ala		
	195	200 205
Gln Arg Ala Ala Ala His Gly Phe Thr Phe Gly Asp Val Arg Thr Thr		
	210	215 220
Phe Thr Gly His Glu Leu Cys Ser Gly Ser Pro Trp Leu His Ser Val		
	225	230 235 240
Asn Trp Leu Asn Ile Gly Glu Ser Tyr His Pro Thr Ala Ala Gly Gln		
	245	250 255
Ser Gly Gly Tyr Leu Pro Val Leu Asn Gly Ala Ala		
	260	265

<210> 32

<211> 269

<212> PRT

<213> Streptomyces avermitilis

<400> 32

Met Arg Arg Ser Arg Ile Thr Ala Tyr Val Thr Ser Leu Leu Leu Ala	
1	5 10 15
Val Gly Cys Ala Leu Thr Gly Ala Ala Thr Ala Gln Ala Ser Pro Ala	
	20 25 30
Ala Ala Ala Thr Gly Tyr Val Ala Leu Gly Asp Ser Tyr Ser Ser Gly	
	35 40 45
Val Gly Ala Gly Ser Tyr Leu Ser Ser Ser Gly Asp Cys Lys Arg Ser	
	50 55 60
Ser Lys Ala Tyr Pro Tyr Leu Trp Gln Ala Ala His Ser Pro Ser Ser	
	65 70 75 80
Phe Ser Phe Met Ala Cys Ser Gly Ala Arg Thr Gly Asp Val Leu Ala	
	85 90 95

115

Asn Gln Leu Gly Thr Leu Asn Ser Ser Thr Gly Leu Val Ser Leu Thr
 100 105 110

Ile Gly Gly Asn Asp Ala Gly Phe Ser Asp Val Met Thr Thr Cys Val
 115 120 125

Leu Gln Ser Asp Ser Ala Cys Leu Ser Arg Ile Asn Thr Ala Lys Ala
 130 135 140

Tyr Val Asp Ser Thr Leu Pro Gly Gln Leu Asp Ser Val Tyr Thr Ala
 145 150 155 160

Ile Ser Thr Lys Ala Pro Ser Ala His Val Ala Val Leu Gly Tyr Pro
 165 170 175

Arg Phe Tyr Lys Leu Gly Gly Ser Cys Leu Ala Gly Leu Ser Glu Thr
 180 185 190

Lys Arg Ser Ala Ile Asn Asp Ala Ala Asp Tyr Leu Asn Ser Ala Ile
 195 200 205

Ala Lys Arg Ala Ala Asp His Gly Phe Thr Phe Gly Asp Val Lys Ser
 210 215 220

Thr Phe Thr Gly His Glu Ile Cys Ser Ser Ser Thr Trp Leu His Ser
 225 230 235 240

Leu Asp Leu Leu Asn Ile Gly Gln Ser Tyr His Pro Thr Ala Ala Gly
 245 250 255

Gln Ser Gly Gly Tyr Leu Pro Val Met Asn Ser Val Ala
 260 265

<210> 33

< 211> 267

< 212> PRT

< 213> Streptomyces sp.

116

<400> 33

Met Arg Leu Thr Arg Ser Leu Ser Ala Ala Ser Val Ile Val Phe Ala
1 5 10 15

Leu Leu Leu Ala Leu Leu Gly Ile Ser Pro Ala Gln Ala Ala Gly Pro
20 25 30

Ala Tyr Val Ala Leu Gly Asp Ser Tyr Ser Ser Gly Asn Gly Ala Gly
35 40 45

Ser Tyr Ile Asp Ser Ser Gly Asp Cys His Arg Ser Asn Asn Ala Tyr

117

50	55	60
Pro Ala Arg Trp Ala Ala Ala Asn Ala Pro Ser Ser Phe Thr Phe Ala		
65	70	75 80
Ala Cys Ser Gly Ala Val Thr Thr Asp Val Ile Asn Asn Gln Leu Gly		
	85	90 95
Ala Leu Asn Ala Ser Thr Gly Leu Val Ser Ile Thr Ile Gly Gly Asn		
	100	105 110
Asp Ala Gly Phe Ala Asp Ala Met Thr Thr Cys Val Thr Ser Ser Asp		
	115	120 125
Ser Thr Cys Leu Asn Arg Leu Ala Thr Ala Thr Asn Tyr Ile Asn Thr		
	130	135 140
Thr Leu Leu Ala Arg Leu Asp Ala Val Tyr Ser Gln Ile Lys Ala Arg		
	145	150 155 160
Ala Pro Asn Ala Arg Val Val Val Leu Gly Tyr Pro Arg Met Tyr Leu		
	165	170 175
Ala Ser Asn Pro Trp Tyr Cys Leu Gly Leu Ser Asn Thr Lys Arg Ala		
	180	185 190
Ala Ile Asn Thr Thr Ala Asp Thr Leu Asn Ser Val Ile Ser Ser Arg		
	195	200 205
Ala Thr Ala His Gly Phe Arg Phe Gly Asp Val Arg Pro Thr Phe Asn		
	210	215 220
Asn His Glu Leu Phe Phe Gly Asn Asp Trp Leu His Ser Leu Thr Leu		
	225	230 235 240
Pro Val Trp Glu Ser Tyr His Pro Thr Ser Thr Gly His Gln Ser Gly		
	245	250 255
Tyr Leu Pro Val Leu Asn Ala Asn Ser Ser Thr		
	260	265

<210> 34

<211> 317

<212> PRT

<213> Aeromonas hydrophila

<400> 34

Ala Asp Ser Arg Pro Ala Phe Ser Arg Ile Val Met Phe Gly Asp Ser
 1 5 10 15

Leu Ser Asp Thr Gly Lys Met Tyr Ser Lys Met Arg Gly Tyr Leu Pro
 20 25 30

Ser Ser Pro Pro Tyr Tyr Glu Gly Arg Phe Ser Asn Gly Pro Val Trp
 35 40 45

Leu Glu Gln Leu Thr Asn Glu Phe Pro Gly Leu Thr Ile Ala Asn Glu
 50 55 60

Ala Glu Gly Gly Pro Thr Ala Val Ala Tyr Asn Lys Ile Ser Trp Asn
 65 70 75 80

Pro Lys Tyr Gln Val Ile Asn Asn Leu Asp Tyr Glu Val Thr Gln Phe
 85 90 95

Leu Gln Lys Asp Ser Phe Lys Pro Asp Asp Leu Val Ile Leu Trp Val
 100 105 110

Gly Ala Asn Asp Tyr Leu Ala Tyr Gly Trp Asn Thr Glu Gln Asp Ala
 115 120 125

Lys Arg Val Arg Asp Ala Ile Ser Asp Ala Ala Asn Arg Met Val Leu
 130 135 140

Asn Gly Ala Lys Glu Ile Leu Leu Phe Asn Leu Pro Asp Leu Gly Gln
 145 150 155 160

Asn Pro Ser Ala Arg Ser Gln Lys Val Val Glu Ala Ala Ser His Val
 165 170 175

Ser Ala Tyr His Asn Gln Leu Leu Leu Asn Leu Ala Arg Gln Leu Ala
 180 185 190

Pro Thr Gly Met Val Lys Leu Phe Glu Ile Asp Lys Gln Phe Ala Glu
 195 200 205

Met Leu Arg Asp Pro Gln Asn Phe Gly Leu Ser Asp Gln Arg Asn Ala
 210 215 220

Cys Tyr Gly Gly Ser Tyr Val Trp Lys Pro Phe Ala Ser Arg Ser Ala
 225 230 235 240

Ser Thr Asp Ser Gln Leu Ser Ala Phe Asn Pro Gln Glu Arg Leu Ala
 245 250 255

Ile Ala Gly Asn Pro Leu Leu Ala Gln Ala Val Ala Ser Pro Met Ala
 260 265 270

Ala Arg Ser Ala Ser Thr Leu Asn Cys Glu Gly Lys Met Phe Trp Asp
275 280 285

Gln Val His Pro Thr Thr Val Val His Ala Ala Leu Ser Glu Pro Ala
290 295 300

Ala Thr Phe Ile Glu Ser Gln Tyr Glu Phe Leu Ala His
305 310 315

- <210> 35
- < 211> 318
- < 212> PRT
- < 213> Aeromonas salmonicida

120

<400> 35

Ala Asp Thr Arg Pro Ala Phe Ser Arg Ile Val Met Phe Gly Asp Ser
 1 5 10 15

Leu Ser Asp Thr Gly Lys Met Tyr Ser Lys Met Arg Gly Tyr Leu Pro
 20 25 30

Ser Ser Pro Pro Tyr Tyr Glu Gly Arg Phe Ser Asn Gly Pro Val Trp
 35 40 45

Leu Glu Gln Leu Thr Lys Gln Phe Pro Gly Leu Thr Ile Ala Asn Glu
 50 55 60

Ala Glu Gly Gly Ala Thr Ala Val Ala Tyr Asn Lys Ile Ser Trp Asn
 65 70 75 80

Pro Lys Tyr Gln Val Ile Asn Asn Leu Asp Tyr Glu Val Thr Gln Phe
 85 90 95

Leu Gln Lys Asp Ser Phe Lys Pro Asp Asp Leu Val Ile Leu Trp Val
 100 105 110

Gly Ala Asn Asp Tyr Leu Ala Tyr Gly Trp Asn Thr Glu Gln Asp Ala
 115 120 125

Lys Arg Val Arg Asp Ala Ile Ser Asp Ala Ala Asn Arg Met Val Leu
 130 135 140

Asn Gly Ala Lys Gln Ile Leu Leu Phe Asn Leu Pro Asp Leu Gly Gln
 145 150 155 160

Asn Pro Ser Ala Arg Ser Gln Lys Val Val Glu Ala Val Ser His Val
 165 170 175

Ser Ala Tyr His Asn Lys Leu Leu Leu Asn Leu Ala Arg Gln Leu Ala
 180 185 190

121

Pro Thr Gly Met Val Lys Leu Phe Glu Ile Asp Lys Gln Phe Ala Glu
 195 200 205

Met Leu Arg Asp Pro Gln Asn Phe Gly Leu Ser Asp Val Glu Asn Pro
 210 215 220

Cys Tyr Asp Gly Gly Tyr Val Trp Lys Pro Phe Ala Thr Arg Ser Val
 225 230 235 240

Ser Thr Asp Arg Gln Leu Ser Ala Phe Ser Pro Gln Glu Arg Leu Ala
 245 250 255

Ile Ala Gly Asn Pro Leu Leu Ala Gln Ala Val Ala Ser Pro Met Ala
 260 265 270

Arg Arg Ser Ala Ser Pro Leu Asn Cys Glu Gly Lys Met Phe Trp Asp
 275 280 285

Gln Val His Pro Thr Thr Val Val His Ala Ala Leu Ser Glu Arg Ala
 290 295 300

Ala Thr Phe Ile Glu Thr Gln Tyr Glu Phe Leu Ala His Gly
 305 310 315

<210> 36

<211> 1371

5

<212> DNA

<213> Streptomyces thermosacchari

<400> 36

```

acaggccgat gcacggaacc gtacctttcc gcagtgaagc gctctccccc catcggtcgc      60
cgggaattca tccgcgattt tggcatgaac acttccttca acgcgcgtag cttgctacaa      120
gtgcggcagc agaccgcctc gttggagget cagtgaagatt gacccgatcc ctgtcggccg      180
catccgtcat cgtcttcgcc ctgctgctcg cgctgctggg catcagcccg gcccaggcag      240
ccggcccggc ctatgtggcc ctgggggatt cctattcctc gggcaacggc gccggaagtt      300
acatcgattc gagcggtgac tgtcacccga gcaacaacgc gtaccccgcc cgctgggcgg      360
cggccaacgc accgtcctcc ttcaccttcg cggcctgctc gggagcggtg accacggatg      420
tgatcaacaa tcagctgggc gccctcaacg cgtccaccgg cctgggtgagc atcaccatcg      480
gcggcaatga cgcgggcttc gcggacgcga tgaccacctg cgtcaccagc tcggacagca      540
cctgcctcaa ccggctggcc accgccacca actacatcaa caccaccctg ctgcgccggc      600
tcgacgcggt ctacagccag atcaaggccc gtgcccccaa cgcgcgcgtg gtcgtcctcg      660
gctaccgcg cgtgtacctg gcctcgaacc cctggtactg cctgggcctg agcaaaccca      720
agcgcgcggc catcaacacc accgcgcgac ccctcaactc ggtgatctcc tcccgggcca      780

ccgcccacgg attccgattc ggcgatgtcc gcccgacctt caacaaccac gaactgttct      840
tcggcaacga ctggctgcac tcaactaccc tgccgggtgtg ggagtcgtac caccaccacca      900
gcacggggcca tcagagcggc tatctgccgg tctcaacgc caacagctcg acctgatcaa      960
cgcacggccg tgcccgcccc gcgcgtcacg ctoggcgccg gcgcgcgagc gcgttgatca     1020
gcccacagtg ccggtgacgg tcccaccgtc acggtcgagg gtgtacgtca cggtggcgcc     1080
gctccagaag tggaacgtca gcaggaccgt ggagccgtcc ctgacctcgt cgaagaactc     1140
cggggtcagc gtgatcacc ctcccccgta gccgggggcg aaggcggcgc cgaactcctt     1200
gtaggacgtc cagtcgtgcg gcccggcgtt gccaccgtcc gcgtagaccg cttccatggt     1260
cgccagccgg tccccgcgga actcggtggg gatgtccgtg cccaaggtgg tcccggtggt     1320
gtccgagagc accggggggt cgtaccggat gatgtgcaga tccaaagaat t              1371

```

5

<210> 37

<211> 267

<212> PRT

<213> Streptomyces thermosacchari

<400> 37

Met Arg Leu Thr Arg Ser Leu Ser Ala Ala Ser Val Ile Val Phe Ala
1 5 10 15

Leu Leu Leu Ala Leu Leu Gly Ile Ser Pro Ala Gln Ala Ala Gly Pro
20 25 30

Ala Tyr Val Ala Leu Gly Asp Ser Tyr Ser Ser Gly Asn Gly Ala Gly
35 40 45

Ser Tyr Ile Asp Ser Ser Gly Asp Cys His Arg Ser Asn Asn Ala Tyr
50 55 60

Pro Ala Arg Trp Ala Ala Ala Asn Ala Pro Ser Ser Phe Thr Phe Ala
65 70 75 80

Ala Cys Ser Gly Ala Val Thr Thr Asp Val Ile Asn Asn Gln Leu Gly
85 90 95

Ala Leu Asn Ala Ser Thr Gly Leu Val Ser Ile Thr Ile Gly Gly Asn
100 105 110

Asp Ala Gly Phe Ala Asp Ala Met Thr Thr Cys Val Thr Ser Ser Asp
115 120 125

Ser Thr Cys Leu Asn Arg Leu Ala Thr Ala Thr Asn Tyr Ile Asn Thr
130 135 140

Thr Leu Leu Ala Arg Leu Asp Ala Val Tyr Ser Gln Ile Lys Ala Arg
145 150 155 160

Ala Pro Asn Ala Arg Val Val Val Leu Gly Tyr Pro Arg Met Tyr Leu
165 170 175

Ala Ser Asn Pro Trp Tyr Cys Leu Gly Leu Ser Asn Thr Lys Arg Ala
180 185 190

Ala Ile Asn Thr Thr Ala Asp Thr Leu Asn Ser Val Ile Ser Ser Arg
195 200 205

Ala Thr Ala His Gly Phe Arg Phe Gly Asp Val Arg Pro Thr Phe Asn
210 215 220

Asn His Glu Leu Phe Phe Gly Asn Asp Trp Leu His Ser Leu Thr Leu
225 230 235 240

Pro Val Trp Glu Ser Tyr His Pro Thr Ser Thr Gly His Gln Ser Gly
245 250 255

Tyr Leu Pro Val Leu Asn Ala Asn Ser Ser Thr
260 265

124

<210> 38

<211> 548

<212> PRT

<213> Thermobifida fusca

5

<400> 38

Met Leu Pro His Pro Ala Gly Glu Arg Gly Glu Val Gly Ala Phe Phe
 1 5 10 15

Ala Leu Leu Val Gly Thr Pro Gln Asp Arg Arg Leu Arg Leu Glu Cys
 20 25 30

His Glu Thr Arg Pro Leu Arg Gly Arg Cys Gly Cys Gly Glu Arg Arg
 35 40 45

Val Pro Pro Leu Thr Leu Pro Gly Asp Gly Val Leu Cys Thr Thr Ser
 50 55 60

Ser Thr Arg Asp Ala Glu Thr Val Trp Arg Lys His Leu Gln Pro Arg
 65 70 75 80

Pro Asp Gly Gly Phe Arg Pro His Leu Gly Val Gly Cys Leu Leu Ala
 85 90 95

Gly Gln Gly Ser Pro Gly Val Leu Trp Cys Gly Arg Glu Gly Cys Arg
 100 105 110

125

Phe Glu Val Cys Arg Arg Asp Thr Pro Gly Leu Ser Arg Thr Arg Asn
 115 120 125

Gly Asp Ser Ser Pro Pro Phe Arg Ala Gly Trp Ser Leu Pro Pro Lys
 130 135 140

Cys Gly Glu Ile Ser Gln Ser Ala Arg Lys Thr Pro Ala Val Pro Arg
 145 150 155 160

Tyr Ser Leu Leu Arg Thr Asp Arg Pro Asp Gly Pro Arg Gly Arg Phe
 165 170 175

Val Gly Ser Gly Pro Arg Ala Ala Thr Arg Arg Arg Leu Phe Leu Gly
 180 185 190

Ile Pro Ala Leu Val Leu Val Thr Ala Leu Thr Leu Val Leu Ala Val
 195 200 205

Pro Thr Gly Arg Glu Thr Leu Trp Arg Met Trp Cys Glu Ala Thr Gln
 210 215 220

Asp Trp Cys Leu Gly Val Pro Val Asp Ser Arg Gly Gln Pro Ala Glu
 225 230 235 240

Asp Gly Glu Phe Leu Leu Leu Ser Pro Val Gln Ala Ala Thr Trp Gly
 245 250 255

Asn Tyr Tyr Ala Leu Gly Asp Ser Tyr Ser Ser Gly Asp Gly Ala Arg
 260 265 270

Asp Tyr Tyr Pro Gly Thr Ala Val Lys Gly Gly Cys Trp Arg Ser Ala
 275 280 285

Asn Ala Tyr Pro Glu Leu Val Ala Glu Ala Tyr Asp Phe Ala Gly His
 290 295 300

Leu Ser Phe Leu Ala Cys Ser Gly Gln Arg Gly Tyr Ala Met Leu Asp
 305 310 315 320

Ala Ile Asp Glu Val Gly Ser Gln Leu Asp Trp Asn Ser Pro His Thr
 325 330 335

Ser Leu Val Thr Ile Gly Ile Gly Gly Asn Asp Leu Gly Phe Ser Thr
 340 345 350

Val Leu Lys Thr Cys Met Val Arg Val Pro Leu Leu Asp Ser Lys Ala
 355 360 365

126

Cys Thr Asp Gln Glu Asp Ala Ile Arg Lys Arg Met Ala Lys Phe Glu
370 375 380

Thr Thr Phe Glu Glu Leu Ile Ser Glu Val Arg Thr Arg Ala Pro Asp
385 390 395 400

Ala Arg Ile Leu Val Val Gly Tyr Pro Arg Ile Phe Pro Glu Glu Pro
405 410 415

Thr Gly Ala Tyr Tyr Thr Leu Thr Ala Ser Asn Gln Arg Trp Leu Asn
420 425 430

Glu Thr Ile Gln Glu Phe Asn Gln Gln Leu Ala Glu Ala Val Ala Val
435 440 445

His Asp Glu Glu Ile Ala Ala Ser Gly Gly Val Gly Ser Val Glu Phe
450 455 460

Val Asp Val Tyr His Ala Leu Asp Gly His Glu Ile Gly Ser Asp Glu
465 470 475 480

Pro Trp Val Asn Gly Val Gln Leu Arg Asp Leu Ala Thr Gly Val Thr
485 490 495

Val Asp Arg Ser Thr Phe His Pro Asn Ala Ala Gly His Arg Ala Val
500 505 510

Gly Glu Arg Val Ile Glu Gln Ile Glu Thr Gly Pro Gly Arg Pro Leu
515 520 525

Tyr Ala Thr Phe Ala Val Val Ala Gly Ala Thr Val Asp Thr Leu Ala
530 535 540

Gly Glu Val Gly
545

<210> 39

<211> 3000

<212> DNA

<213> Thermobifida fusca

<400> 39

ggtggtgaac cagaacaccc ggtcgtcggc gtgggcgtcc aggtgcaggt gcaggttctt	60
caactgctcc agcaggatgc cggcgtggcc gtgcacgatg gccttgggca ggcctgtggt	120
ccccgacgag tacagcacc atagcggatg gtogaacggc agcgggggtga actccagttc	180
cgcgccttcg cccgcggcctt cgaactccgc ccaggacagg gtgtcggcga cagggccgca	240
gcccaggtac ggcaggacga cgggtgtgctg caggctgggc atgccgtcgc gcagggcttt	300

gagcacgtca cggcgggtcga agtccttacc gccgtagcgg tagccgtcca cggccagcag	360
cactttcgggt tcgatctgcg cgaaccggtc gaggacgctg cgcaccccga agtcggggga	420
acaggacgac caggtcgcac cgatcgcggc gcaggcgagg aatgcggccg tcgcctcggc	480
gatgttcggc aggtaggcca cgaaccggtc gccggggccc accccgaggc tgcggagggc	540
cgcagcgatc gcggcgggtgc gggtcgcgag ttctccccag gtccactcgg tcaacggccg	600
gagttcggac gcgtgccgga tcgccacggc tgatgggtca cggtcgcgga agatgtgctc	660
ggcgtagttg aggggtggcg cggggaacca gacggcgccg ggcatggcgt cggaggcga	720
cactgtgtg tacgggttg cggcgcgac ccggtagtag tcccagatcg cggaccagaa	780
tccttcgagg tcggttaccg accagcgcca cagtgcctcg tagtccggtg cgtccacacc	840
gcgggtgctc cgcaccacgc gggtagacgc ggtgaggttg gcgcgttctt tgcgtcctc	900
gtcgggactc cacaggatcg gcggctcggc cttgagtgtc atgaaacgcg accccttcgt	960
ggacgggtcgc gatgcggtga gcgtcgggtg cctcccctaa cgctcccgg tgacggagt	1020
ttgtgcacca catctagcac gcgggacgcg gaaaccgtat ggagaaaaca cctacaaccc	1080
cggccggacg gtgggtttcg gccacactta ggggtcgggt gcctgcttgc cgggcagggc	1140
agtcccgggg tgctgtggtg cgggcgggag ggctgtcgt tcgagggtgt cggcgggac	1200
actccgggcc tcagccgtac ccgcaacggg gacagtctc ctcccttcg ggcgtggtg	1260
tccttcccc cgaaatgcg cgagatctc cagtacgccc ggaacacac cgctgtgcc	1320
aggtactctt tgcttcgaac agacaggccg gacgggtccac gggggagggt tgtgggcagc	1380
ggaccacgtg cggcgaccag acgacggtt ttctcggta tcccgcctct tgtacttgt	1440
acagcgctca cgctggtctt ggctgtccg acggggcgcg agacgctgtg gcgcattgtg	1500
tgtgaggcca ccaggactg gtgcctgggg gtgcgggtcg actcccgcg acagcctgcg	1560
gaggacggcg agtttctgct gctttctccg gtccaggcag cgacctgggg gaactattac	1620
gcgctcgggg attcgtactc ttccggggac ggggcccgcg actactatcc cggcacccgcg	1680
gtgaaggcg gttgctggcg gtccgctaac gcctatccg agctggtcgc cgaagcctac	1740
gaattcgcg gacacttgct gttoctggcc tgcagcgcc agcgcggcta cgcctgctt	1800
gacgctatcg acgaggtcgg ctgcagctg gactggaact cccctcacac gtcgctggtg	1860
acgatcggga tcggcggcaa cgatctgggg ttctccacgg ttttgaagac ctgcatggtg	1920
cgggtgcgc tgctggacag caaggcgtgc acggaccagg aggacgctat ccgcaagcgg	1980
atggcgaaat tcgagacgac gtttgaagag ctcatcagcg aagtgcgcac ccgcgcgcg	2040
gacgcccggg tccttgctgt gggctacccc cggatcttcc cggaggaaac gaccggcgcc	2100
tactacacgc tgaccgcgag caaccagcgg tggctcaacg aaaccattca ggagttcaac	2160
cagcagctcg ccgaggtgt cgcgggtccac gacgaggaga ttgcgcgctc gggcgggggtg	2220
ggcagcgtgg agttcgtgga cgtctaccac gcgttgacg gccacgagat cggctcggac	2280

gagccgtggg tgaacggggt gcagttgcgg gacctcgcca cgggggtgac tgtggaccgc 2340
 agtacettcc accccaacgc cgctgggcac cgggcggtcg gtgagcgggt catcgagcag 2400
 atcgaaaccg gcccgggccg tccgctctat gccactttcg cgggtggtggc gggggcgacc 2460
 gtggacactc tcgcgggcga ggtggggtga cccggcctac cgtccggccc gcaggtctgc 2520
 gagcactgcg gcgatctggt ccaactgccc gtgcagttcg tcttcggtga tgaccagcgg 2580
 cggggagagc cggatcggtg agccgtgcgt gtctttgacg agcacacccc gctgcaggag 2640
 ccgttcgcac agttctcttc cgggtggccag agtcgggtcg acgtcgatcc cagcccacag 2700
 gccgatgctg cgggcgcgga ccacgcggtt gccgaccagt tggtcgaggc ggcgcgcgag 2760
 cacgggggag agggcgcgga catggtccag gtaagggccg tcgcggacga ggctcaccac 2820
 ggcagtgcgc accgcgcagg cgaggcggtt gccgcggaag gtgctgccgt gctggccggg 2880
 gcggatcacg tcgaagactt ccgcgtcgcc taaccgcgcc gccacgggca ggatgccgcc 2940
 gccacgcgct ttgccgaaca ggtagatata ggcgctgact ccgctgtggt cgcaggcccg 3000

<210> 40

<211> 372

<212> PRT

<213> Thermobifida fusca

<400> 40

Val Gly Ser Gly Pro Arg Ala Ala Thr Arg Arg Arg Leu Phe Leu Gly
 1 5 10 15
 Ile Pro Ala Leu Val Leu Val Thr Ala Leu Thr Leu Val Leu Ala Val
 20 25 30
 Pro Thr Gly Arg Glu Thr Leu Trp Arg Met Trp Cys Glu Ala Thr Gln
 35 40 45
 Asp Trp Cys Leu Gly Val Pro Val Asp Ser Arg Gly Gln Pro Ala Glu
 50 55 60
 Asp Gly Glu Phe Leu Leu Leu Ser Pro Val Gln Ala Ala Thr Trp Gly
 65 70 75 80
 Asn Tyr Tyr Ala Leu Gly Asp Ser Tyr Ser Ser Gly Asp Gly Ala Arg
 85 90 95
 Asp Tyr Tyr Pro Gly Thr Ala Val Lys Gly Gly Cys Trp Arg Ser Ala
 100 105 110
 Asn Ala Tyr Pro Glu Leu Val Ala Glu Ala Tyr Asp Phe Ala Gly His
 115 120 125

Leu Ser Phe Leu Ala Cys Ser Gly Gln Arg Gly Tyr Ala Met Leu Asp
 130 135 140

Ala Ile Asp Glu Val Gly Ser Gln Leu Asp Trp Asn Ser Pro His Thr
 145 150 155 160

Ser Leu Val Thr Ile Gly Ile Gly Gly Asn Asp Leu Gly Phe Ser Thr
 165 170 175

Val Leu Lys Thr Cys Met Val Arg Val Pro Leu Leu Asp Ser Lys Ala
 180 185 190

Cys Thr Asp Gln Glu Asp Ala Ile Arg Lys Arg Met Ala Lys Phe Glu
 195 200 205

Thr Thr Phe Glu Glu Leu Ile Ser Glu Val Arg Thr Arg Ala Pro Asp
 210 215 220

Ala Arg Ile Leu Val Val Gly Tyr Pro Arg Ile Phe Pro Glu Glu Pro
 225 230 235 240

Thr Gly Ala Tyr Tyr Thr Leu Thr Ala Ser Asn Gln Arg Trp Leu Asn
 245 250 255

Glu Thr Ile Gln Glu Phe Asn Gln Gln Leu Ala Glu Ala Val Ala Val
 260 265 270

His Asp Glu Glu Ile Ala Ala Ser Gly Gly Val Gly Ser Val Glu Phe
 275 280 285

Val Asp Val Tyr His Ala Leu Asp Gly His Glu Ile Gly Ser Asp Glu
 290 295 300

Pro Trp Val Asn Gly Val Gln Leu Arg Asp Leu Ala Thr Gly Val Thr
 305 310 315 320

Val Asp Arg Ser Thr Phe His Pro Asn Ala Ala Gly His Arg Ala Val
 325 330 335

Gly Glu Arg Val Ile Glu Gln Ile Glu Thr Gly Pro Gly Arg Pro Leu
 340 345 350

Tyr Ala Thr Phe Ala Val Val Ala Gly Ala Thr Val Asp Thr Leu Ala
 355 360 365

Gly Glu Val Gly
 370

<210> 41

<211> 300

130

< 212> PRT

< 213> Corynebacterium efficiens

<400> 41

Met Arg Thr Thr Val Ile Ala Ala Ser Ala Leu Leu Leu Leu Ala Gly
 1 5 10 15

Cys Ala Asp Gly Ala Arg Glu Glu Thr Ala Gly Ala Pro Pro Gly Glu
 20 25 30

Ser Ser Gly Gly Ile Arg Glu Glu Gly Ala Glu Ala Ser Thr Ser Ile
 35 40 45

Thr Asp Val Tyr Ile Ala Leu Gly Asp Ser Tyr Ala Ala Met Gly Gly
 50 55 60

Arg Asp Gln Pro Leu Arg Gly Glu Pro Phe Cys Leu Arg Ser Ser Gly
 65 70 75 80

Asn Tyr Pro Glu Leu Leu His Ala Glu Val Thr Asp Leu Thr Cys Gln
 85 90 95

Gly Ala Val Thr Gly Asp Leu Leu Glu Pro Arg Thr Leu Gly Glu Arg
 100 105 110

Thr Leu Pro Ala Gln Val Asp Ala Leu Thr Glu Asp Thr Thr Leu Val
 115 120 125

Thr Leu Ser Ile Gly Gly Asn Asp Leu Gly Phe Gly Glu Val Ala Gly
 130 135 140

Cys Ile Arg Glu Arg Ile Ala Gly Glu Asn Ala Asp Asp Cys Val Asp
 145 150 155 160

Leu Leu Gly Glu Thr Ile Gly Glu Gln Leu Asp Gln Leu Pro Pro Gln
 165 170 175

Leu Asp Arg Val His Glu Ala Ile Arg Asp Arg Ala Gly Asp Ala Gln
 180 185 190

Val Val Val Thr Gly Tyr Leu Pro Leu Val Ser Ala Gly Asp Cys Pro
 195 200 205

Glu Leu Gly Asp Val Ser Glu Ala Asp Arg Arg Trp Ala Val Glu Leu
 210 215 220

Thr Gly Gln Ile Asn Glu Thr Val Arg Glu Ala Ala Glu Arg His Asp
 225 230 235 240

131

Ala Leu Phe Val Leu Pro Asp Asp Ala Asp Glu His Thr Ser Cys Ala
245 250 255

Pro Pro Gln Gln Arg Trp Ala Asp Ile Gln Gly Gln Gln Thr Asp Ala
260 265 270

Tyr Pro Leu His Pro Thr Ser Ala Gly His Glu Ala Met Ala Ala Ala
275 280 285

Val	Arg	Asp	Ala	Leu	Gly	Leu	Glu	Pro	Val	Gln	Pro
290						295					300

<210> 42

 $\langle 211 \rangle$ 3000

5 < 212> DNA

< 213> *Corynebacterium efficiens*

<400> 42

ttctgggggtg ttatgggggtt gttatcggct cgtcctgggt ggatcccgcc aggtggggta	60
ttcacggggg acttttgtgt ccaacagccg agaatgagt ccctgagcgg tgggaatgag	120
gtgggcgggg ctgtgtcgcc atgagggggc ggcgggctct gtggtgcccc gcgacccccg	180
gccccggtga gcggtgaatg aaatccggct gtaatcagca tcccgtagcc acccgtcgg	240
ggaggtcagc gcccgagtg tctacgcagt cggatcctct cggactcggc catgctgtcg	300
gcagcatcgc gctcccgggt cttggcgcc ctcggctgtt ctgcctgctg tccttggaag	360
gcgaaatgat caccggggag tgatacaccg gtggtctcat ccgggatgcc cacttcggcg	420
ccatccggca attcgggcag ctccgggtgg aagtaggtgg catccgatgc gtcggtgacg	480
ccatagtggg cgaagatctc atcctgctcg aggggtgctca ggcactctc cggatcgata	540
tggggggcgt ccttgatggc gtccttgctg aaaccgaggt gcagcttgctg ggcttccaat	600
ttcgaccac ggagcgggac gaggtggaa tgacggccga agagcccggt gtggacctca	660
acgaaggtgg gtagtcccggt gtcattcattg aggaacacgc cctccaccgc acccagcttg	720
tggccggagt tgctgtaggc gctggcatcc agaagggaaa cgatctcata tttgtcgggtg	780
tgctcagaca tgatcttctt ttgctgtcgg tgtctgttac taccacggtg gggtgaatg	840
caactgttat ttttctgtta ttttaggaat tggccatat cccacaggct ggctgtggtc	900
aaatcgtcat caagtaatcc ctgtcacaca aaatgggtgg tgggagccct ggtcgcggtt	960
ccgtgggagg cgcctgccc cgcaggatcg tcggcatcgg cggatctggc cggtagccccg	1020
cggtagaataa aatcattctg taaccttcac caccgttggt tttaggtatc cggcccttcc	1080
gtcctgaccc cgtccccggc gcgcgggagc ccgcgggttg cggtagacag gggagacgtg	1140
gacaccatga ggacaaaggc catcgcagca agcgcattac tccttctcgc cggatgcgcg	1200
gatggggccc gggaggagac cgcgggtgca ccgcgggtg agtcctccgg gggcatccgg	1260

gaggaggggg cggaggcgtc gacaagcatc accgacgtct acatcgccct cggggattcc 1320
 tatgcggcga tgggcgggcg ggatcagccg ttacgggggtg agccgttctg cctgcgctcg 1380
 tccggttaatt acccggaact cctccacgca gaggtcacccg atctcacctg ccagggggcg 1440
 gtgaccgggg atctgctcga acccaggacg ctggggggagc gcacgctgcc ggcgcagggtg 1500
 gatgcgctga cggaggacac caccctggtc accctctcca tcgggggcaa tgacctcgga 1560
 ttcggggagg tggcgggatg catccgggaa cggatcgccg gggagaacgc tgatgattgc 1620
 gtggacctgc tgggggaaac catcggggag cagctcgatc agcttcccc gcagctggac 1680
 cgcgtgcacg aggtctatccg ggaccgcgcg ggggacgcgc aggttggtgt caccggttac 1740
 ctgccgctcg tgtctgccgg ggactgcccc gaactggggg atgtctccga ggcggatcgt 1800
 cgttggggcg ttgagctgac cgggcagatc aaogagaccg tgcgcgaggc ggccgaacga 1860
 cacgatgccc tctttgtcct gcccgcacgat gccgatgagc acaccagttg tgcaccccca 1920
 cagcagcgct gggcggatat ccagggccaa cagaccgatg cctatccgct gcacccgacc 1980
 tccgcgggcc atgaggcgat ggccgccgcc gtccgggacg cgctgggcct ggaaccggtc 2040
 cagccgtagc gccgggcgcg cgcttgctga cgaccaaccc atgccaggct gcagtcacat 2100
 ccgcacatag cgcgcgcggg cgatggagta cgcaccatag aggatgagcc cgatgccgac 2160
 gatgatgagc agcacactgc cgaagggttg tccccgagg gtgcgcagag ccgagtccag 2220
 acctgcggcc tgctccggat catgggcccc accggcgatg acgatcaaca cccccaggat 2280
 cccgaaggcg ataccacggg cgacataacc ggctgttccg gtgatgatga tcgcgggtccc 2340
 gacctgccct gaccccgcac ccgcctccag atcctccccg aaatccccgg tggccccctt 2400
 ccagagggtg tagacacccg cccccagtac caccagcccg gcgaccacaa ccagcaccac 2460
 accccagggt tgggatagga cggtgccggt gacatcggtg gcggtctccc catcggaggt 2520
 gctgccgccc cgggcgaagg tggaggtggt caccgccagg gagaagtaga ccatggccat 2580
 gaccgcccc ttggccottt ccttgaggtc ctgcgccgc agcagctggc tcaattgcca 2640
 gagtcccgagg gccgccaggg cgatgacggc aaccacagg aggaactgcc caccgggagc 2700
 ctccgcgatg gtggccaggg cacctgaatt cgaggcctca tcaccgaac cgccggatcc 2760
 agtggcgatg cgcaccgca tccaccgat gaggatgtgc agtatgcca ggacaatgaa 2820
 accacotctg gccagggtgg tcagcgcggg gtggtcctcg gcctggtcgg cagcccgctc 2880
 gatcgtccgt ttcgcggatc tgggtgcgcc cttatccata gctccattg aacgccttg 2940
 aggggtgggc ggccactgtc agggcggatt gtgatctgaa ctgtgatgtt ccatcaaccc 3000

<210> 43

<211> 268

<212> PRT

<213> Streptomyces coelicolor

<400> 43

```

Met Arg Arg Phe Arg Leu Val Gly Phe Leu Ser Ser Leu Val Leu Ala
1      5      10      15

Ala Gly Ala Ala Leu Thr Gly Ala Ala Thr Ala Gln Ala Ala Gln Pro
      20      25      30

Ala Ala Ala Asp Gly Tyr Val Ala Leu Gly Asp Ser Tyr Ser Ser Gly
      35      40      45

Val Gly Ala Gly Ser Tyr Ile Ser Ser Ser Gly Asp Cys Lys Arg Ser
      50      55      60

Thr Lys Ala His Pro Tyr Leu Trp Ala Ala Ala His Ser Pro Ser Thr
      65      70      75      80

Phe Asp Phe Thr Ala Cys Ser Gly Ala Arg Thr Gly Asp Val Leu Ser
      85      90      95

Gly Gln Leu Gly Pro Leu Ser Ser Gly Thr Gly Leu Val Ser Ile Ser
      100      105      110

Ile Gly Gly Asn Asp Ala Gly Phe Ala Asp Thr Met Thr Thr Cys Val
      115      120      125

Leu Gln Ser Glu Ser Ser Cys Leu Ser Arg Ile Ala Thr Ala Glu Ala
      130      135      140

Tyr Val Asp Ser Thr Leu Pro Gly Lys Leu Asp Gly Val Tyr Ser Ala
      145      150      155      160

Ile Ser Asp Lys Ala Pro Asn Ala His Val Val Val Ile Gly Tyr Pro
      165      170      175

Arg Phe Tyr Lys Leu Gly Thr Thr Cys Ile Gly Leu Ser Glu Thr Lys
      180      185      190

Arg Thr Ala Ile Asn Lys Ala Ser Asp His Leu Asn Thr Val Leu Ala
      195      200      205

Gln Arg Ala Ala Ala His Gly Phe Thr Phe Gly Asp Val Arg Thr Thr
      210      215      220

Phe Thr Gly His Glu Leu Cys Ser Gly Ser Pro Trp Leu His Ser Val
      225      230      235      240

Asn Trp Leu Asn Ile Gly Glu Ser Tyr His Pro Thr Ala Ala Gly Gln
      245      250      255

Ser Gly Gly Tyr Leu Pro Val Leu Asn Gly Ala Ala
      260      265

```

<210> 44

< 211> 2000

< 212> DNA

< 213> Streptomyces coelicolor

5

<400> 44

```

cccggcggcc cgtgcaggag cagcagccgg cccgcgatgt cctcgggcgt cgtcttcate      60
aggccgtcca tcgcgtcgcc gaccggcgcc gtgtagttag cccggacctc gtcccagggtg      120
cccgcggcga tctggcgggt ggtgcgggtgc gggccgcgcc gaggggagac gtaccagaag      180
cccatcgta cgttctccgg ctgcggttcg ggctcgtaac ccgctccgtc cgtcgccctcg      240
ccgagcacct tctcggcgag gtcggcgctg gtcgccgtca ccgtgacgtc ggcgccccgg      300
ctccagcgcg agatcagcag cgtccagccg tcgccctccg ccagcgtagc gctcgggtcg      360
tcgtcggggg cgatccgcag cagcgcgcg cccggcgcca gcagcgtagc gccggaccgt      420
acgcgggtcga tgttcgcggc gtgcgagtag ggctgctcac ccgtggcgaa acggccgagg      480
aacagcgcggt cgacgacgtc ggacggggag tcgctgtcgt ccacgttgag ccggatcggc      540
agggcttcgt gcgggttcac ggacatgtcg ccgatgtagc gcacccggcc gccgcgtgca      600
cccgttttcc cgggcacgca cgacaggggc tttctcgccg tcttcctgcc gaacttgaac      660
gagtgtcagc cattttcttg catggacact tccagtcaac gcgcgttagc gctaccacgg      720
ttgtggcagc aatcctgcta agggagggtc catgagacgt ttccgacttg tcggcttcct      780
gagttcgctc gtctcgcggc cgggcgcggc cctcaccggg gcagcgaccg cccaggcggc      840
ccaacccgcc gcgcgcgacg gctatgtggc cctcggcgac tctactcct cccgggtcgg      900
agcgggcagc tacatcagct cgagcgcgga ctgcaagcgc agcacgaagg cccatcccta      960
cctgtggggc gccgcccact cgccctccac gttcgacttc accgcctgtt ccggcgcccg      1020
tacgggtgat gttctctccg gacagctcgg cccgctcagc tccggcaccg gcctcgtctc      1080
gatcagcadc ggcggcaacg acgcccgttt cgccgacacc atgacgacct gtgtgctcca      1140
gtccgagagc tctgcctgt cgcggtatcg caccgcggag gcgtacgtcg actcgacgtc      1200
gcccggcaag ctcgacggcg tctactcggc aatcagcgac aaggcgccga acgcccacgt      1260
cgtcgtcadc ggctacccgc gcttctacaa gctcggcaac acctgcatcg gcctgtccga      1320
gaccaagcgg acggcgatca acaaggcctc cgaccacctc aacaccgtcc tcgcccagcg      1380
cgccgcggcc cagcggttca ccttcgggga cgtacgcacc accttcaccg gccacgagct      1440
gtgctccggc agcccctggc tgcacagcgt caactggctg aacatcggcg agtcgtacca      1500
ccccaccggc gccggccagt ccggtggcta cctgccggtc ctcaacggcg ccgctgacc      1560
tcaggcgga gagaagaag aaggagcgga gggagacgag gaggggagg ccccgcccga      1620
cggggtcccc gtccccgtct ccgtctccgt cccggtcccg caagtcaccg agaacgccac      1680

```

```

cgcgtcggac gtggcccgca cgggactccg cacctccacg cgcacggcac tctcgaacgc      1740
gccgggtgtcg tcgtgcgtcg tcaccaccac gccgtcctgg cgcgagcgct cgcgcgccga      1800
cgggaaggac agcgtccgcc accccggatc ggagaccgac ccgtccgcgg tcaccaccg      1860
gtagccgacc tccgcgggca gccgcccgc cgtgaacgtc gccgtgaacg cgggtgcccg      1920
gtcgtgctggc ggcggacagg cccccgagta gtgggtgcgc gagccacca cggtcacctc      1980
caccgactgc gctgcggggc                                2000

```

<210> 45

<211> 269

5

<212> PRT

<213> Streptomyces avermitilis

<400> 45

```

Met Arg Arg Ser Arg Ile Thr Ala Tyr Val Thr Ser Leu Leu Leu Ala
1           5           10           15

Val Gly Cys Ala Leu Thr Gly Ala Ala Thr Ala Gln Ala Ser Pro Ala
                20           25           30

Ala Ala Ala Thr Gly Tyr Val Ala Leu Gly Asp Ser Tyr Ser Ser Gly
          35           40           45

Val Gly Ala Gly Ser Tyr Leu Ser Ser Ser Gly Asp Cys Lys Arg Ser
50           55           60

Ser Lys Ala Tyr Pro Tyr Leu Trp Gln Ala Ala His Ser Pro Ser Ser
65           70           75           80

Phe Ser Phe Met Ala Cys Ser Gly Ala Arg Thr Gly Asp Val Leu Ala
          85           90           95

Asn Gln Leu Gly Thr Leu Asn Ser Ser Thr Gly Leu Val Ser Leu Thr
100          105          110

Ile Gly Gly Asn Asp Ala Gly Phe Ser Asp Val Met Thr Thr Cys Val
115          120          125

Leu Gln Ser Asp Ser Ala Cys Leu Ser Arg Ile Asn Thr Ala Lys Ala
130          135          140

Tyr Val Asp Ser Thr Leu Pro Gly Gln Leu Asp Ser Val Tyr Thr Ala
145          150          155          160

Ile Ser Thr Lys Ala Pro Ser Ala His Val Ala Val Leu Gly Tyr Pro
165          170          175

```

137

Arg Phe Tyr Lys Leu Gly Gly Ser Cys Leu Ala Gly Leu Ser Glu Thr
 180 185 190

Lys Arg Ser Ala Ile Asn Asp Ala Ala Asp Tyr Leu Asn Ser Ala Ile
 195 200 205

Ala Lys Arg Ala Ala Asp His Gly Phe Thr Phe Gly Asp Val Lys Ser
 210 215 220

Thr Phe Thr Gly His Glu Ile Cys Ser Ser Ser Thr Trp Leu His Ser
 225 230 235 240

Leu Asp Leu Leu Asn Ile Gly Gln Ser Tyr His Pro Thr Ala Ala Gly
 245 250 255

Gln Ser Gly Gly Tyr Leu Pro Val Met Asn Ser Val Ala
 260 265

<210> 46

<211> 1980

5

<212> DNA

<213> Streptomyces avermitilis

<400> 46

ccaccgcgcg	gtcggcgggcg	agtctcctgg	cctcggtcgc	ggagagggtg	gccgtgtagc	60
cgttcagcgc	ggcgccgaac	gtcttcttca	ccgtgccgcc	gtactcgttg	atcaggccct	120
tgcccttgct	cgacgcggcc	ttgaagccgg	tgcccttctt	gagcgtgacg	atgtagctgc	180
ccttgatcgc	ggtggggggag	ccggcgggga	gcaccgtgcc	ctcggccggg	gtggcctggg	240
cgggcagtcg	ggtgaatccg	cccacgaggg	cgccggtcgc	cacggcggtt	atcgccggca	300
tccggtctt	cttgctacgc	agctgtgcca	tacgagggag	tcctcctctg	ggcagcggcg	360
cgcctgggtg	gggcgcacgg	ctgtgggggg	tgcgcgcgtc	atcacgcaca	cggccctgga	420
gcgtcgtgtt	ccgccctggg	ttgagtaaag	cctcggccat	ctacgggggt	ggctcaaggg	480
agttgagacc	ctgtcatgag	tctgacatga	gcacgcaatc	aacggggccg	tgagcacccc	540
ggggcgaccc	cggaaagtgc	cgagaagtct	tggcatggac	acttctctgc	aacacgcgta	600
gctggtagca	cggttacggc	agagatcctg	ctaaagggag	gttccatgag	acgttccoga	660
attaoggcct	acgtgacctc	actcctcctc	gccgtcggtt	gcgccctcac	cggggcagcg	720
acggcgcagg	cgtccccagc	cgccgcggcc	acgggctatg	tggccctcgg	cgactcgtac	780
tgcctcgggtg	tggcgccggg	cagctaacctc	agctccagcg	gcgactgcaa	gcgcagttcg	840
aaggcctatc	cgtacctctg	gcaggccggc	cattcacccct	cgtcgttcag	tttcatggct	900
tgcctcggggc	ctcgtacggg	tgatgtcctg	gccaatcagc	tcggcaccct	gaactcgtcc	960
accggcctgg	tctccctcac	catcgagggc	aacgacgcgg	gcttctccga	cgtcatgacg	1020
acctgtgtgc	tccagtccga	cagcgccctgc	ctctcccgca	tcaacacggc	gaaggcgtac	1080
gtcgactcca	ccctgcccgg	ccaactcgac	agcgtgtaca	cggcgatcag	cacgaaggcc	1140
ccgtcggccc	atgtggccgt	gctgggctac	cccgccttct	acaaactggg	cggctcctgc	1200
ctcgcggggc	tctcggagac	caagcggtcc	gccatcaacg	acgcggccga	ctatctgaac	1260
agcggccatcg	ccaagcgcgc	cgccgaccac	ggcttcacct	tcggcgacgt	caagagcacc	1320
ttcacccggc	atgagatctg	ctccagcagc	acctggctgc	acagtctcga	cctgctgaac	1380
atcggccagt	cctaccaccc	gaccgcggcc	ggccagtcgg	gcggctatct	gccggtcatg	1440
aacagcgtgg	cctgagctcc	cacggcctga	atctttaagg	cctgaatttt	taaggcgaag	1500
gtgaaccgga	agcggaggcc	ccgtccgtcg	gggtctccgt	cgcacaggtc	accgagaacg	1560
gcacggagtt	ggacgtcgtg	cgcaccgggt	cgcgcacctc	gacggcgatc	tcgttcgaga	1620
tcgttccgct	cgtgtcgtac	gtggtgacga	acacctgctt	ctgctgggtc	tttccgccgc	1680
tcggccgggaa	ggacagcgtc	ttccagcccg	gatccgggac	ctcgccttct	ttggtcacc	1740
agcgggtactc	cacctcgacc	ggcaccgggc	ccaccgtgaa	ggtcggccgtg	aacgtggggc	1800
cctggggcgtg	ggggcggggg	caggcaccgg	agtagtccgt	gtgcacgcgc	gtgaccgtca	1860
ccttcacgga	ctgggcccggc	ggggtcgtcg	taccgcggcc	gccaccggcc	cctcccgag	1920
tggagcccga	gctgtgggtcg	ccccgcctgt	cggcggtgtc	gtcctcgggg	gttttcgaac	1980

139

<210> 47

<211> 372

<212> PRT

<213> Thermobifida fusca

5

<400> 47

Met	Gly	Ser	Gly	Pro	Arg	Ala	Ala	Thr	Arg	Arg	Arg	Leu	Phe	Leu	Gly
1				5					10					15	

Ile	Pro	Ala	Leu	Val	Leu	Val	Thr	Ala	Leu	Thr	Leu	Val	Leu	Ala	Val
			20					25					30		

Pro	Thr	Gly	Arg	Glu	Thr	Leu	Trp	Arg	Met	Trp	Cys	Glu	Ala	Thr	Gln
		35					40					45			

Asp	Trp	Cys	Leu	Gly	Val	Pro	Val	Asp	Ser	Arg	Gly	Gln	Pro	Ala	Glu
	50					55					60				

Asp	Gly	Glu	Phe	Leu	Leu	Leu	Ser	Pro	Val	Gln	Ala	Ala	Thr	Trp	Gly
65					70					75					80

Asn	Tyr	Tyr	Ala	Leu	Gly	Asp	Ser	Tyr	Ser	Ser	Gly	Asp	Gly	Ala	Arg
				85					90					95	

Asp	Tyr	Tyr	Pro	Gly	Thr	Ala	Val	Lys	Gly	Gly	Cys	Trp	Arg	Ser	Ala
			100					105					110		

140

Asn Ala Tyr Pro Glu Leu Val Ala Glu Ala Tyr Asp Phe Ala Gly His
115 120 125

Leu Ser Phe Leu Ala Cys Ser Gly Gln Arg Gly Tyr Ala Met Leu Asp
130 135 140

Ala Ile Asp Glu Val Gly Ser Gln Leu Asp Trp Asn Ser Pro His Thr
145 150 155 160

Ser Leu Val Thr Ile Gly Ile Gly Gly Asn Asp Leu Gly Phe Ser Thr
165 170 175

Val Leu Lys Thr Cys Met Val Arg Val Pro Leu Leu Asp Ser Lys Ala
180 185 190

Cys Thr Asp Gln Glu Asp Ala Ile Arg Lys Arg Met Ala Lys Phe Glu
195 200 205

Thr Thr Phe Glu Glu Leu Ile Ser Glu Val Arg Thr Arg Ala Pro Asp
210 215 220

Ala Arg Ile Leu Val Val Gly Tyr Pro Arg Ile Phe Pro Glu Glu Pro
225 230 235 240

Thr Gly Ala Tyr Tyr Thr Leu Thr Ala Ser Asn Gln Arg Trp Leu Asn
245 250 255

Glu Thr Ile Gln Glu Phe Asn Gln Gln Leu Ala Glu Ala Val Ala Val
260 265 270

His Asp Glu Glu Ile Ala Ala Ser Gly Gly Val Gly Ser Val Glu Phe
275 280 285

Val Asp Val Tyr His Ala Leu Asp Gly His Glu Ile Gly Ser Asp Glu
290 295 300

Pro Trp Val Asn Gly Val Gln Leu Arg Asp Leu Ala Thr Gly Val Thr
305 310 315 320

Val Asp Arg Ser Thr Phe His Pro Asn Ala Ala Gly His Arg Ala Val
325 330 335

Gly Glu Arg Val Ile Glu Gln Ile Glu Thr Gly Pro Gly Arg Pro Leu
340 345 350

Tyr Ala Thr Phe Ala Val Val Ala Gly Ala Thr Val Asp Thr Leu Ala
355 360 365

Gly Glu Val Gly
370

<210> 48

<211> 968

< 212> DNA

< 213> Thermobifida fusca

<400> 48

ctgcagacac ccgccccgcc ttctcccgga tcgtcatggt cggcgactcc ctcagcgaca	60
ccggcaagat gtactccaag atgcgcggtt acctgccgtc ctccccgcgc tactacgagg	120
gccgtttctc gaacggcccg gtctggctgg agcagctgac gaagcagttc cccggcctga	180
cgatcgccaa cgagggcgag gggggcgaga ccgcagtcgc ctacaacaag atctcctgga	240
acccgaagta ccaggtcatt aacaacctcg actacgaggt caccagttc ttgcagaagg	300
actcgttcaa gcccgacgac ctgggtcatcc tgtgggtggg cgccaacgac tacctggcct	360
acggttggaa caccgagcag gacgccaagc ggggtgcgga cgccatctcg gacggggcaa	420
accgcatggt cctgaacggc gcgaagcaga tcctgctgtt caacctgccg gacctgggcc	480
agaaccgctc cgcccgctcc cagaaggctc tcgaggccgt ctgcgacgtg tccgcctacc	540
acaacaagct gctcctcaac ctgcgccggc agctcgcccc gacgggcatg gtcaagctgt	600
tcgagatcga caagcagttc ggggagatgc tgcgcgaccc ccagaacttc ggctgagcg	660
acgtggagaa cccgtgctac gacggcggtt acgtgtggaa gccgttcgcc acccggtccg	720
tctcgaccga ccggcagctg tcggccttct cgccccagga gcgcctggcg atcgctggca	780
accgctcctt ggcacaggcg gtagcttcgc cgatggcccg ccgctcggcc tcgccccca	840
actgcgaggg caagatgttc tgggaccagg tccacccac caccgtggtc caccgcgcc	900
tctcgagagc cgccgccacc ttcctcgaga ccagtagca gttcctcgcc cactagtcta	960
gaggatcc	968

5

<210> 49

< 211> 1044

< 212> DNA

< 213> Aeromonas salmonicida

<400> 49

atgaacaac aaaaacggct ttacgccga ttgctgacgc tggtatttgc gtcctcttc	60
ttgctgcctc attctgcagc ttcagcagca gatacaagac cggcgtttag ccggatcgtc	120
atgtttggag atagcctgag cgatacgggc aaaatgtata gcaaatgag aggctatctt	180
ccgtcaagcc cgccgtatta tgaaggccgc tttagcaatg gaccggtctg gctggaacaa	240
ctgacgaaac aatttccggg actgacgac gctaataag cagaaggagg agcaacagcg	300
gtcgccctata acaaaatcag ctgggacccg aaatatcagg tcatcaacaa cctggactat	360
gaagtcacac agtttcttca gaaagacagc tttaaacggg atgatctggt catcctttgg	420

10

```

gtcggcgcca atgattatct ggcgatggc tggacacag aacaagatgc caaaagagtc 480
agagatgcca tcagcgatgc cgctaataga atggctctga acggcgccaa acaaatcctg 540
ctgtttaacc tgccggatct gggacaaaat ccgagcgcca gaagccaaaa agtcgtcgaa 600
gcagtcagcc atgtcagcgc ctatcataac aaactgctgc tgaacctggc aagacaattg 660
gcaccgacgg gaatgggttaa attgtttgaa attgacaaac agtttgccga aatgctgaga 720
gatccgcaaa attttggcct gagcgatgtc gaaaaccctg gctatgatgg cggatatgtc 780
tggaaaccgt ttgccacaag aagcgtcagc acggatagac aactgtcagc gtttagcccg 840
caagaagac tggcaatcgc cggaaatccg cttttggcac aagcagttgc ttcaccgatg 900
gcaagaagat cagcaagccc gctgaattgc gaaggcaaaa tgttttggga tcaggtccat 960
ccgacaacag ttgtccatgc tgccctttca gaaagagcgg cgacgtttat cgaaacacag 1020
tatgaatttc tggcccatgg ctga 1044

```

<210> 50

<211> 1005

<212> DNA

<213> *Aeromonas hydrophila*

<400> 50

```

atgaaaaaat ggtttgtgtg tttattggga ttggctcgcg tgacagttca ggcagccgac 60
agccgtcccc ccttctcccg gatcgtgatg tttggcgaca gcctctccga taccggcaag 120
atgtacagca agatgcgcgg ttacctcccc tccagcccc cctactatga gggcgccttc 180
tccaacgggg ccgtctggct ggagcagctg accaacgagt tccggggcct gaccatagcc 240
aacgaggcgg aaggcggacc gaccgcccgt gcttacaaca agatctcctg gaatcccaag 300
tatcaggtca tcaacaacct ggaactagag gtcaccaggt tcctgcaaaa agacagcttc 360
aagccggacg atctggtgat cctctgggtc ggcgccaacg actatctggc ctatggctgg 420
aacacagagc aggatgccaa gcgggtgcgc gacgccatca gcgatgcggc caaccgcatg 480
gtgctgaacg gcgccaagga gatactgctg ttcaacctgc cggatctggg ccagaacccc 540
tcggccccga gccagaaggt ggtcgaggcg gccagccatg tctccgcta ccacaaccag 600
ctgctgctga acctggcacg ccagctggct cccaccggca tggatgaagct gttcgagatc 660
gacaagcagt ttgccgagat gctgcgtgat ccgcagaact tcggcctgag cgaccagagg 720
aacgcctgct acggtggcag ctatgtatgg aagccgtttg cctccgcag cgcagcacc 780
gacagccagc tctccgcctt caaccgcag gagcgccctg ccacgcggg caaccgctg 840
ctggccdagc ccgtcgccag ccccatggct gccgcagcg ccagcacctt caactgtgag 900
ggcaagatgt tctgggatca ggtccacccc accactgtcg tgcaagccgc cctgagcgag 960
ccgcgcgcca ccttcacga gagccagtac gagttcctcg cccac 1005

```

<210> 51

<211> 1011

<212> DNA

<213> *Aeromonas salmonicida*

5

<400> 51

```

atgaaaaaat ggtttgtttg tttattgggg ttgatcgcgc tgacagttca ggcagccgac      60
actcgccccc ccttctcccg gatcgtgatg ttgggcgaca gcctctccga taccggcaaa      120
atgtacagca agatgcgcgg ttacctcccc tccagcccg cctactatga gggccgtttc      180
tccaacggac ccgtctggct ggagcagctg accaagcagt tccgggtct gaccatcgcc      240
aacgaagcgg aaggcgggtgc cactgccgtg gcttacaaca agatctcctg gaatcccaag      300
tatcaggtct acaacaacct ggactacgag gtcacccagt tcttgagaa agacagcttc      360
aagccggacg atctggtgat cctctgggtc ggtgccaatg actatctggc atatggctgg      420
aatacggagc aggatgccaa gcgagttcgc gatgccatca gcgatgcggc caaccgcatg      480
gtactgaacg gtgccaaagca gatactgctg ttcaacctgc cggatctggg ccagaaccgg      540
tcagcccgca gtcagaaggt ggtcgaggcg gtcagccatg tctccgccta tcacaacaag      600
ctgctgctga acctggcacg ccagctggcc cccaccggca tggtaaagct gttcgagatc      660
gacaagcaat ttgccgagat gctgcgtgat ccgcagaact tcggcctgag cgacgtcgag      720
aaccctgtct acgacggcgg ctatgtgtgg aagccgtttg ccaccgcag cgtcagcacc      780
gaccgccagc tctccgcctt cagtccgcag gaagccctcg ccatcgccgg caaccgctg      840
ctggcacagg ccgttgccag tctatggcc cgcgcgagcg ccagccccc caactgtgag      900
ggcaagatgt tctgggatca ggtacacccg accactgtcg tgcacgcagc cctgagcgag      960
cgcgcgcgca ccttcacga gaccagtac gagttcctcg cccacggatg a              1011

```

<210> 52

<211> 888

<212> DNA

10

<213> *Streptomyces coelicolor*

<400> 52

atgccgaagc ctgcccttcg ccgtgtcatg accgcgacag tcgccgccgt cggcacgctc	60
gccctcggcc tcaccgacgc caccgcccac gccgcgcccg cccaggccac tccgacctg	120
gactacgtcg ccctcggcga cagctacagc gccggctccg gcgtcctgcc cgtcgacccc	180
gccaaectgc tctgtctgcg ctcgacggcc aactaccccc acgtcatcgc ggacacgacg	240
ggcgcccgcc tcacggacgt cacctgcggc gccgcgcaga ccgccgactt caccggggcc	300
cagtaccogg gcgtcgcacc ccagttggac gcgtcggca ccggcacgga cctggtcacg	360
ctcaccatcg gcggcaacga caacagcacc ttcatacaag ccatacaggc ctgcggcacg	420
gcgggtgtcc tcagcggcgg caagggcagc ccctgcaagg acaggcacgg cacctccttc	480
gacgacgaga tcgaggccaa cactacccc gcgtcaagg aggcgtgct cggcgtccgc	540
gccagggctc cccacgccag ggtggcggct ctgggtacc cgtggatcac cccggccacc	600
gccgaccogt cctgcttcct gaagctcccc ctgcgcgccg gtgacgtgcc ctacctgcgg	660
gccatccagg cacacctcaa cgacgcggtc cggcgggccg ccgaggagac cggagccacc	720
tacgtggact tctccggggt gtccgacggc cacgacgcct gcgaggcccc cggcaccgcg	780
tggatcgaac cgctgctctt cgggcacagc ctggtcccg tccacccaa cgcctgggc	840
gagcggcgca tggccgagca cacgatggac gtctcggcc tggactga	888

5

<210> 53

<211> 888

<212> DNA

<213> Streptomyces coelicolor

<400> 53

tcagtcacagg ccgaggacgt ccatcgtgtg ctccggccatg cgcgcgtcgc ccagggcggt	60
gggggtggacg ggaacgaggg tgtgcccgaa gagcagcggg tcgatccagc ggggtgccggg	120
ggcctcgcag gcgtcgtggc cgtcggacac cccggagaag tccacgtagg tggtccggg	180
ctcctcggcg gcccgccgga ccgcgtcgtt gaggtgtgcc tggatggccc gcaggtaggg	240
cacgtcaccg gcggcgaggg ggagcttcag gaagcaggac gggtcggcgg tgccgggggt	300
gatccacggg tagccgagag ccgccaccct ggcgtgggga gccctggcgc ggacgccgag	360
cagcgccctc ttgagcgcg ggtacgtgtt ggcctcgatc tcgtcgtcga aggaggtgcc	420
gtgcctgtcc ttgcaggggc tgcccttgcc gccgctgagg acaccgccg tgccgcaggc	480
cgtgatggcg ttgatgaagg tgctgttgtc gttgccgccg atggtgagcg tgaccaggtc	540
cgtgcgggtg ccgagcgcgt ccaactgggg tgdcagcccc gggtaactggg cccgcgtgaa	600
gtcggcggtc tgccggcgcc ccgaggtgac gtccgtgagg cgggcgcccc tcgtgtccgc	660
gatgacgtgg gggtagttgg ccgtcagcgc cagacagagc aggttggcgg ggtcgacggg	720
caggacgccg gagccggcgc tgtagctgtc gccgagggcg acgtagtcca gggtcggagt	780
ggcctgggcg ggcgcggcgt gggcggtggc gtcggtgagg ccgagggcga gcgtgccgac	840
ggcggcgact gtcgcggtea tgacacggcg aagggcaggc ttcggcat	888

<210> 54

<211> 717

5

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 54

atggattacg agaagtttct gttatttggg gattccatta ctgaatttgc ttttaatact	60
aggccattg aagatggcaa agatcagtat gctcttgagg ccgcattagt caacgaatat	120
acgagaaaaa tggatattct tcaaagaggg ttcaaagggt acacttctag atgggcgttg	180
aaaatacttc ctgagatttt aaagcatgaa tccaatattg tcatggccac aatatttttg	240
ggcgccaacg atgcatgctc agcaggtccc caaagtgtcc ccctccccga atttatcgat	300
aatatctgct aaatggtatc tttgatgaag tcttaccata tccgtcctat tataatagga	360
ccggggctag tagatagaga gaagtgggaa aaagaaaaat ctgaagaaat agctctcgga	420
tacttccgta ccaacgagaa ctttgccatt tattccgatg ccttagcaaa actagccaat	480
gaggaaaaag ttcccttcgt ggctttgaat aaggcgtttc aacaggaagg tggatgatgt	540
tggcaacaac tgctaacaga tggactgcac ttttccggaa aagggtacaa aatttttcat	600
gacgaattat tgaaggatcat tgagacattc taccaccaat atcatcccaa aaacatgcag	660
tacaaactga aagattggag agatgtgcta gatgatggat ctaacataat gtcttga	717

10

<210> 55

< 211> 1044

< 212> DNA

< 213> *Ralstonia* sp.

5

<400> 55

```

atgaacctgc gtcaatggat gggcgccgcc acggetgccc ttgccttggg cttggccgcg      60
tgcggggggcg gtgggaccga ccagagcggc aatcccaatg tcgccaaggt gcagcgcag      120
gtggtgttcg gcgacagcct gagcgatata ggcacctaca cccccgtcgc gcaggcggcg      180
ggcggcggca agttcaccac caaccggggc ccgatctggg ccgagaccgt ggccgcgcaa      240
ctgggcgtga cgctcacgcc ggcggtgatg ggctacgcca cctccgtgca gaattgcccc      300
aaggccggct gcttcgacta tgcgcagggc ggctcgcgcg tgaccgatcc gaacggcatc      360
ggccacaacg gcggcgcggg ggcgctgacc tacccggttc agcagcagct cgccaacttc      420
tacgcggcca gcaacaacac attcaacggc aataacgatg tcgtcttcgt gctggccggc      480
agcaacgaca ttttcttctg gaccactcgc gcggccacca gcggtcccg cgtgacgccc      540
gccattgcca cggcccaggt gcagcaggcc gcgacggacc tggtcggcta tgtcaaggac      600
atgatcgcca aggggtcgcg gcaggtctac gtgttcaacc tgcccgcagc cagcctgacg      660
ccggacggcg tggcaagcgg cacgacgggc caggcgctgc tgcaacgcgt ggtgggcacg      720
ttcaacacga cgctgcaaag cgggctggcc ggcacctcgg cgcgcacat caacttcaac      780
gcacaactga ccgcgcgcat ccagaatggc gcctcgttcg gcttcgcca caccagcgcc      840
cgggcctcgc acgccacca gatcaatgcc ctggtgcoga gcgcggcgcg cagctcgcg      900
ttctgctcgg ccaacacgct ggtggcttcc ggtgcggacc agagctacct gttcgcgcg      960
ggcggtgcac cgaccacggc cggccatcgc ctgatcgcca gcaacgtgct ggcgcgcctg     1020
ctggcggata acgtcgcgca ctga                                     1044

```

<210> 56

< 211> 786

< 212> DNA

10

< 213> *Streptomyces coelicolor*

<400> 56

gtgatcgggt cgtacgtggc ggtgggggac agcttcaccg agggcgctcg cgaccccggc	60
cccgaacggg cgttcgtcgg ctgggcccgc cggctcgccg tactgctcgc ggaccggcgc	120
cccgaaggcg acttcacgta cacgaacctc gccgtgcgcg gcaggctcct cgaccagatc	180
gtggcggaac aggtcccgcg ggtcgtcggc ctgcgcgccg acctcgtctc gttcgcggcg	240
ggcggcaacg acatcatccg gcccggcacc gatcccgacg aggtcgccga gcggttcgag	300
ctggcggtgg ccgcgctgac cgcgcggccc ggaaccgtcc tggtgaccac cgggttcgac	360
acccgggggg tgcccgtcct caagcacctg cgcggcaaga tcgccacgta caacgggcac	420
gtcccgccca tcgccgaccg ctacggctgc ccggtgctcg acctgtggtc gctgcggagc	480
gtccaggacc gcaggggcgt ggacgcccgc cggctgcacc tgtcgccgga ggggcacacc	540
cgggtggcgc tcgcgcgggg gcaggccctg ggcctgcgcg tcccggccga ccctgaccag	600
ccctggccgc ccctgcgcgc gcgcggcacg ctgcacgtcc ggcgcgacga cgtgcactgg	660
gcgcgcgagt acctggtgcc gtggatcggg ccgcggctgc ggggcgagtc gtcggggcac	720
cacgtgacgg ccaaggggac gctgtcgccg gacgccatca agacgcggat cgcgcgggtg	780
gcctga	786

<210> 57

<211> 783

<212> DNA

<213> Streptomyces coelicolor

<400> 57

atgcagacga accccgcgta caccagtctc gtgcgcgtcg gcgactcctt caccgagggc	60
atgtcggacc tgctgcccga cggctcctac cgtggctggg ccgacctcct cggcaccggg	120
atggcggccc gctcccccgg cttccggtac gccaacctgg cggtgccgcg gaagctgac	180
ggacagatcg tcgacgagca ggtggacgtg gccgcgccca tgggagccga cgtgatcacg	240
ctggtcggcg ggctcaacga cacgtgcgg cccaagtgcg acatggcccg ggtgcgggac	300
ctgctgaccc aggcctgga acggctcgcc ccgactgcg agcagctggt gctgatgcgc	360
agtcccggtc gccagggtcc ggtgctggag cgttcccggc cccgcatgga ggccctgttc	420
gccgtgatcg acgacctggc cgggcggcac ggcgcctgg tgcgcgacct gtacggggcc	480
cagtcgctgg ccgacctcg gatgtgggac gtggaccggc tgcacctgac cgcgagggc	540
caccgccggg tcgcggaggc ggtgtggcag tcgctcgccc acgagcccga ggaccccgag	600
tggcacgcgc cgatcccggc gacgcggccc cgggggtggg tgacgcgcag gaccgcggac	660
gtccggttcg cccggcagca cctgctgccc tggataggcc gcaggctgac cgggcgctcg	720
tccggggacg gcctgccggc caagcggccc gacctgctgc cctacgagga ccccgcacgg	780
tga	783

<210> 58

<211> 1365

<212> DNA

<213> Streptomyces coelicolor

5

<400> 58

```

atgacccggg gtcgtgacgg ggggtgcggg gcgcccccca ccaagcaccg tgccctgctc      60
gcgggcgatcg tcaccctgat agtggcgatc tcgcgggcca tatacgccgg agcgtccgcg      120
gacgacggca gcagggaacca cgcgctgcag gccggaggcc gtctcccacg aggagacgcc      180
gcccccgct ccaccgggtgc ctgggtgggc gcctgggcca ccgcaccggc cgcggccgag      240
ccgggcaccg agacgaccgg cctggcgggc cgctccgtgc gcaacgtcgt gcacacctcg      300
gtcggcgggca ccggcgcgcg gatcacctc tcgaacctgt acgggcagtc gccgctgacc      360
gtcacacacg cctcgatcgc cctggccgcc gggcccgaca ccgcgcgcgc gatcgccgac      420
accatgcgcc ggctcaoctt cggcggcagc gcccggtga tcatcccgcc gggcgggccag      480
gtgatgagcg acaccgccc cctcgccatc ccctacgggg cgaacgtcct ggtcaccacg      540
tactccccc tcccgctccg gccggtgacc taccatccgc aggcccgcca gaccagctac      600
ctggccgacg gcgaccgcac ggcggacgtc accgccgtcg cgtacaccac cccacgccc      660
tactggcgct acctgaccgc cctcgacgtg ctgagccacg aggcggacgg cacggtcgtg      720
gcggttcggcg actccatcac cgacggcgcc cgctcgacga gcgacgcca ccaccgctgg      780
accgacgtcc tcgcccacag cctgcacgag gcggcgggcg accggccggga cgcgccccgc      840
tacagcgtcg tcaacgaggg catcagcgcc aaccggctcc tgaccagcag gccggggcgg      900
ccggccgaca acccgagcgg actgagccgg ttccagcggg acgtgctgga acgcaccaac      960
gtcaaggccg tcgtcgtcgt cctcgcgctc aacgacgtcc tgaacagccc ggaactcgcc      1020
gaccgcgacg ccatactgac cggcctgcgc accctcgtcg accggcgcca cgcgcgggga      1080
ctgcgggtcg tcggcgccac gatcacgccc ttcgggcggt acggcggcta caccgaggcc      1140
cgcgagacga tgcggcagga ggtcaacgag gagatccgct ccggccgggt cttcgacacg      1200
gtcgtcgact tcgacaaggc cctgcgcgac ccgtacgacc cgcgcgggat gcgctccgac      1260
tacgacagcg gcgaccacct gcaccccgcc gacaaggggt acgcgcgcat gggcgcggtc      1320
atcgacctgg ccgcgctgaa gggcgcgggc ccggtcaagg cgtag      1365

```

<210> 59

<211> 1023

<212> DNA

<213> Streptomyces coelicolor

10

<400> 59

atgacgagca tgtcgagggc gaggggtggcg cggcgggatcg cggccggcgc ggcgtacggc	60
ggcggcggca tcggcctggc gggagcggcg gcggtcggtc tgggtggcgc cgaggtgcag	120
ctggccagac gcaggggtggg ggtgggcacg ccgacccggg tgcggaacgc gcagggactg	180
tacggcggca ccctgcccac ggccggcgac ccgccgctgc ggctgatgat gctgggcgac	240
tccacggccg ccgggcaggc cgtgcaccgg gccgggcaga cgcggggcgc gctgctggcg	300
tccgggtcgc cggcgggtggc ggagcggccg gtgcggctgg ggtcggtcgc ccagccgggg	360
gcgtgctcgc acgacctgga ccggcaggct gcgctggtgc tcgccgagcc ggaccgggtg	420
cccacatct gcgtgatcat ggtcggcgcc aacgacgtca cccaccgat gccggcgacc	480
cgtcgggtgc ggcacctgtc ctccggcgta cggcggctgc gcacggccgg tcgggaggtg	540
gtggtcggca cctgtccgga cctgggcacg atcgagcggg tcgggcagcc gctgcgctgg	600
ctggcccgcc gggcctcacg gcagctcgcg gcggcacaga ccacggcgcc cgtcgagcag	660
ggcgggcgca cgggtgcgct gggcgacctg ctgggtccgg agttcgcgca gaaccgcgg	720
gagctcttcg gcccgcacaa ctaccacccc tcggccgagg ggtacgccac ggccgcgatg	780
gcggtactgc cctcgggtgtg cgcgcgcgtc ggcctgtggc cggccgacga ggagcaccgc	840
gacgcgctgc gccgcgaggc cttcctgccg gtggcgcgcg cggcggcgga ggcggcgtcc	900
gaggcgggta cggaggtcgc cgcgcccatg cctacggggc ctcgggggcc ctgggcgctg	960
ctgaagcgcc ggagacggcg tcgggtgtcg gaggcggaac cgtccagccc gtccggcggtt	1020
tga	1023

5

<210> 60

<211> 918

<212> DNA

<213> Streptomyces coelicolor

<400> 60

atgggtcgag ggacggacca ggggacggg tacggccgtc gccggggcgg tgtcgcgcto	60
gccgccctga ccgccgccgt cctgggcggtg ggcgtggcgg gctgcgactc cgtgggcggc	120
gactcaccgg ctccctccgg cagcccgtcg aagcggacga ggacggcgcc cgcctgggac	180
accagcccgg cgtccgtcgc cggcgtgggc gactccatca cgcgcggctt cgacgcctgt	240
gcggtgctgt cggactgcc ggaggtgtcg tgggcgaccg gcagcagcgc gaaggtcgac	300
tcgctggcgg tacggtgct ggggaaggcg gacgcggcgg agcacagctg gaactacgcg	360
gtcaccgggg cccggtatgg ggacctgacc gctcaggtga cgcgggcggc gcagcgcgag	420
ccggaagctg tggcggtgat ggccggggcg aacgacgcgt gccggtccac gacctcggcg	480
atgacgccgg tggcggaact cggggcgcg ttogaggagg cgatggccac cctgcgcaag	540
aagctcccca aggcgcaggt gtacgtgtcg agcatcccg acctcaagcg gctctggtcc	600
cagggccgca ccaaccgct gggcaagcag gtgtggaagc tcggcctgtg cccgtcgatg	660
ctggggcgac cggactccct ggactcggcg gcgaccctgc ggcgcaacac ggtgcgcgac	720
cgggtggcgg actacaacga ggtgctgcgg gaggtctcgg cgaaggaccg gcggtgccgc	780
agcgacgacg gcgcggtgca cgagttccgg ttcggcacgg accagttgag ccaactgggac	840
tggttccacc cgagtgtgga cggccaggcc cggctggcgg agatcgcta ccgcgcggtc	900
 accgcgaaga atccctga	 918

5

<210> 61

<211> 1068

<212> DNA

<213> Streptomyces rimosus

<400> 61

ttcatcacaa cgatgtcaca acaccggcca tccgggtcat ccctgatcgt gggaatgggt	60
gacaagcctt cccgtgacga aagggtcctg ctacatcaga aatgacagaa atcctgctca	120
gggaggttcc atgagactgt cccgacgcgc ggccacggcg tccgcgctcc tcctacccc	180
ggcgctcgcg ctcttcggcg cgagcgccgc cgtgtccgcg ccgcgaatcc aggccaccga	240
ctacgtggcc ctggcgact cctactcctc gggggtcggc gcgggcagct acgacagcag	300
cagtggctcc tgtaagcgca gcaccaagtc ctaccgggc ctgtgggccc cctcgcacac	360
cggtagcgcg ttcaacttca ccgcctgttc gggcgcccgc acaggagacg tgctggccaa	420
gcagctgacc ccggtcaact ccggcaccga cctggtcagc attaccatcg gcggcaacga	480
cgcgggcttc gccgacacca tgaccacctg caacctccag ggcgagagcg cgtgcctggc	540
gcggatcgcc aaggcgcgcg cctacatcca gcagacgctg cccgcccagc tggaccaggt	600
ctacgacgcc atcgacagcc gggccccgc agcccaggtc gtcgtcctgg gctacccgcg	660
cttctacaag ctggggcgga gctgcgcgt cgggtctctcg gagaagtccc gcgcggccat	720
caacgccgcc gccgacgaca tcaacgcgt caccgccaag cgcgccgcg accacggctt	780
cgccttcggg gacgtcaaca cgaccttcgc cgggcacgag ctgtgctccg gcgccccctg	840
gctgcacagc gtcacccttc ccgtggagaa ctccctaccac cccacggcca acggacagtc	900
caagggtac ctgcccgctc tgaactcgc cactgatct cgcggtact ccgccctga	960
cgaagtccc ccccgggcg gggcttcgcc gtaggtgcgc gtaccgcgt cgcccgtcgc	1020
gccggtggcc ccgcgtagc tgccgccgc cccggacgc gtcggttc	1068

<210> 62

<211> 1008

<212> DNA

<213> *Aeromonas hydrophila*

<400> 62

atgaaaaaat ggtttgtgtg ttattggga ttggtcgcgc tgacagttca ggcagccgac	60
agtcgcccc ccttttccc gatcgtgatg ttggcgaca gcctctccga tacggcaaa	120
atgtacagca agatgcggcg ttacctccc tccagccgc cctactatga gggcgttcc	180
tccaacggac ccgtctggct ggagcagctg accaaacagt tcccgggtct gaccatcgcc	240
aacgaagcgg aaggcggtgc cactgccgtg gcttacaaca agatctcctg gaatcccaag	300
tatcaggtca tcaacaacct ggactacgag gtcaccagc tttgcagaa agacagcttc	360

```

aagccggacg atctggtgat cctctgggtc ggtgccaatg actatctggc ctatggctgg 420
aacacggagc aggatgccaa gcgggttcgc gatgccatca gcgatgcggc caaccgcatg 480
gtactgaacg gtgccaagca gatactgctg ttcaacctgc cggatctggg ccagaaccgg 540
tcagctcgca gtcagaaggt ggtcgaggcg gtcagccatg tctccgccta tcacaaccag 600
ctgctgctga acctggcacg ccagctggcc cccaccggca tggtaaagct gttcgagatc 660
gacaagcaat ttgccgagat gctgcgtgat ccgcagaact tcggcctgag cgacgtcgag 720
aaccctgct acgacggcgg ctatgtgtgg aagccgtttg ccaccgcag cgtcagcacc 780
gaccgccagc tctccgcctt cagtcgcag gaacgcctcg ccacgcggc caaccgctg 840
ctggcacagg ccgttgccag tcctatggcc cgcgcgagcg ccagccccct caactgtgag 900
ggcaagatgt tctgggatca ggtacacccg accactgtcg tgcacgcagc cctgagcgag 960
cgcgcgcgca ccttcacgc gaaccagtac gagttcctcg cccactga 1008

```

<210> 63

<211> 1011

<212> DNA

<213> *Aeromonas salmonicida*

<400> 63

```

atgaaaaaat ggtttgtttg ttattgggg ttgatcgcg tgacagttca ggcagccgac 60
actcgcccc ccttctccc gatcgtgat ttggcgaca gcctctcca taccggcaaa 120
atgtacagca agatgcggg ttacctccc tcacgccgc cctactatga gggccgtttc 180
tccaacggac ccgtctggct ggagcagctg accaagcagt tccgggtct gaccatcgcc 240
aacgaagcgg aaggcgggtgc cactgcctg gcttacaaca agatctctg gaatcccaag 300
tatcaggtca tcaacaacct ggactacgag gtcaccagc tcttgagaa agacagcttc 360
aagccggacg atctggtgat cctctgggtc ggtgccaatg actatctggc atatggctgg 420
aatacggagc aggatgccaa gcgagttcgc gatgccatca gcgatgcggc caaccgcatg 480
gtactgaacg gtgccaagca gatactgctg ttcaacctgc cggatctggg ccagaaccgg 540
tcagcccgca gtcagaaggt ggtcgaggcg gtcagccatg tctccgccta tcacaacaag 600
ctgctgctga acctggcacg ccagctggcc cccaccggca tggtaaagct gttcgagatc 660
gacaagcaat ttgccgagat gctgcgtgat ccgcagaact tcggcctgag cgacgtcgag 720
aaccctgct acgacggcgg ctatgtgtgg aagccgtttg ccaccgcag cgtcagcacc 780
gaccgccagc tctccgcctt cagtcgcag gaacgcctcg ccacgcggc caaccgctg 840
ctggcacagg ccgttgccag tcctatggcc cgcgcgagcg ccagccccct caactgtgag 900
ggcaagatgt tctgggatca ggtacacccg accactgtcg tgcacgcagc cctgagcgag 960
cgcgcgcgca ccttcacga gaccagtac gagttcctcg cccacggatg a 1011

```

153

<210> 64

< 211> 51

< 212> DNA

< 213> Artificial

5

<220>

< 223> Terminator sequence

<400> 64

cgggacttac cgaaagaac catcaatgat gggttctttt ttgtcataa a 51

<210> 65

10

< 211> 59

< 212> DNA

< 213> Artificial

<220>

< 223> Terminator sequence

15

<400> 65

caagactaaa gaccgttcgc ccgttttgc aataagcggg cgaatcttac ataaaaata 59

<210> 66

< 211> 61

< 212> DNA

20

< 213> Artificial

<220>

< 223> Terminator sequence

<400> 66

acggccggtta gatgtgacag ccggttccaa aaggaagcgg gdtgtcttcg tgtattattg 60

t 61

25

<210> 67

< 211> 54

< 212> DNA

< 213> Artificial

154

<220>

< 223> Terminator sequence

<400> 67

tcttttaaag gaaaggctgg aatgcccggc attccagcca catgatcatc gttt 54

5

<210> 68

< 211> 56

< 212> DNA

< 213> Artificial

<220>

10

< 223> Terminator sequence

<400> 68

gctgacaaat aaaaagaagc aqgtatggag gaacctgctt cttttacta ttattg 56

<210> 69

< 211> 102

15

< 212> DNA

< 213> Artificial

<220>

< 223> Oligonucleotide primer

<400> 69

ccccgctcga ggcttttctt ttggaagaaa atatagggaa aatgggtactt gttaaaaatt 60

20

cggaatatatt atacaatatc atatgtttca cattgaaagg gg 102

<210> 70

< 211> 35

< 212> DNA

< 213> Artificial

25

<220>

< 223> Oligonucleotide primer

155

<400> 70

tggaatctcg aggtttatc cttaccttg tctcc

35

<210> 71

 $\langle 211 \rangle 8$

5

< 212> PRT

< 213> Artificial

 $\langle 220 \rangle$

< 223> Consensus sequence

<400> 71

Met Arg Arg Ser Arg Phe Leu Ala
1 5

10

<210> 72

 $\langle 211 \rangle 8$

< 212> PRT

< 213> Artificial

15

 $\langle 220 \rangle$

< 223> Consensus sequence

<400> 72

Ala Leu Ile Leu Leu Thr Leu Ala
1 5

20

<210> 73

 $\langle 211 \rangle 5$

< 212> PRT

< 213> Artificial

 $\langle 220 \rangle$

< 223> Consensus sequence

25

<400> 73

Ala Arg Ala Ala Pro
1 5

156

<210> 74

< 211> 11

< 212> PRT

< 213> Artificial

5

<220>

< 223> Consensus sequence

<400> 74

Tyr	Val	Ala	Leu	Gly	Asp	Ser	Tyr	Ser	Ser	Gly
1				5					10	

<210> 75

10

< 211> 5

< 212> PRT

< 213> Artificial

<220>

< 223> Consensus sequence

15

<400> 75

Gly	Ala	Gly	Ser	Tyr
1			5	

<210> 76

< 211> 4

< 212> PRT

20

< 213> Artificial

<220>

< 223> Consensus sequence

<400> 76

Ser	Ser	Gly	Asp
1			

25

<210> 77

< 211> 15

157

< 212> PRT

< 213> Artificial

<220>

< 223> Consensus sequence

5 <400> 77

Arg	Ser	Thr	Lys	Ala	Tyr	Pro	Ala	Leu	Trp	Ala	Ala	Ala	His	Ala
1				5					10				15	

<210> 78

< 211> 5

< 212> PRT

10 < 213> Artificial

<220>

< 223> Consensus sequence

<400> 78

Ser	Ser	Phe	Ser	Phe
1			5	

15 <210> 79

< 211> 12

< 212> PRT

< 213> Artificial

<220>

20 < 223> Consensus sequence

<400> 79

Ala	Cys	Ser	Gly	Ala	Arg	Thr	Tyr	Asp	Val	Leu	Ala
1				5					10		

<210> 80

< 211> 15

25 < 212> PRT

< 213> Artificial

158

<220>

< 223> Consensus sequence

<400> 80

Leu	Val	Ser	Ile	Thr	Ile	Gly	Gly	Asn	Asp	Ala	Gly	Phe	Ala	Asp
1				5				10						15

5

<210> 81

< 211> 6

< 212> PRT

< 213> Artificial

<220>

10

< 223> Consensus sequence

<400> 81

Met	Thr	Thr	Cys	Val	Leu
1				5	

<210> 82

< 211> 6

15

< 212> PRT

< 213> Artificial

<220>

< 223> Consensus sequence

<400> 82

20

Ser	Asp	Ser	Ala	Cys	Leu
1				5	

<210> 83

< 211> 4

< 212> PRT

< 213> Artificial

25

<220>

< 223> Consensus sequence

159

<400> 83

Thr Leu Pro Ala
1

<210> 84

< 211> 9

5 < 212> PRT

< 213> Artificial

<220>

< 223> Consensus sequence

<400> 84

Arg Leu Asp Ser Val Tyr Ser Ala Ile
1 5

<210> 85

< 211> 4

< 212> PRT

< 213> Artificial

15 <220>

< 223> Consensus sequence

<400> 85

Thr Arg Ala Pro
1

<210> 86

20 < 211> 12

< 212> PRT

< 213> Artificial

<220>

< 223> Consensus sequence

25 <400> 86

Ala Arg Val Val Val Leu Gly Tyr Pro Arg Ile Tyr
1 5 10

160

<210> 87

< 211> 4

< 212> PRT

< 213> Artificial

5

<220>

< 223> Consensus sequence

<400> 87

Leu Gly Leu Ser
1

<210> 88

10

< 211> 11

< 212> PRT

< 213> Artificial

<220>

< 223> Consensus sequence

15

<400> 88

Thr Lys Arg Ala Ala Ile Asn Asp Ala Ala Asp
1 5 10

<210> 89

< 211> 12

< 212> PRT

20

< 213> Artificial

<220>

< 223> Consensus sequence

<400> 89

Leu Asn Ser Val Ile Ala Lys Arg Ala Ala Asp His
1 5 10

25

<210> 90

< 211> 7

161

< 212> PRT

< 213> Artificial

<220>

< 223> Consensus sequence

5 <400> 90

Gly	Phe	Thr	Phe	Gly	Asp	Val
1				5		

<210> 91

< 211> 7

< 212> PRT

10 < 213> Artificial

<220>

< 223> Consensus sequence

<400> 91

Gly	His	Glu	Leu	Cys	Ser	Ala
1				5		

15 <210> 92

< 211> 9

< 212> PRT

< 213> Artificial

<220>

20 < 223> Consensus sequence

<400> 92

Pro	Trp	Leu	His	Ser	Leu	Thr	Leu	Pro
1				5				

<210> 93

< 211> 6

25 < 212> PRT

< 213> Artificial

162

<220>

< 223> Consensus sequence

<400> 93

Ser Tyr His Pro Thr Ala
 1 5

5

<210> 94

< 211> 13

< 212> PRT

< 213> Artificial

<220>

10

< 223> Consensus sequence

<400> 94

Gly His Ala Ala Gly Tyr Leu Pro Val Leu Asn Ser Ile
 1 5 10

<210> 95

< 211> 232

15

< 212> PRT

< 213> *Aspergillus aculeatus*

<400> 95

Thr Thr Val Tyr Leu Ala Gly Asp Ser Thr Met Ala Lys Asn Gly Gly
 1 5 10 15

Gly Ser Gly Thr Asn Gly Trp Gly Glu Tyr Leu Ala Ser Tyr Leu Ser
 20 25 30

Ala Thr Val Val Asn Asp Ala Val Ala Gly Arg Ser Ala Arg Ser Tyr
 35 40 45

Thr Arg Glu Gly Arg Phe Glu Asn Ile Ala Asp Val Val Thr Ala Gly
 50 55 60

Asp Tyr Val Ile Val Glu Phe Gly His Asn Asp Gly Gly Ser Leu Ser
 65 70 75 80

Thr Asp Asn Gly Arg Thr Asp Cys Ser Gly Thr Gly Ala Glu Val Cys
 85 90 95

Tyr Ser Val Tyr Asp Gly Val Asn Glu Thr Ile Leu Thr Phe Pro Ala

163

```

          100              105              110

Tyr Leu Glu Asn Ala Ala Lys Leu Phe Thr Ala Lys Gly Ala Lys Val
    115              120              125

Ile Leu Ser Ser Gln Thr Pro Asn Asn Pro Trp Glu Thr Gly Thr Phe
    130              135              140

Val Asn Ser Pro Thr Arg Phe Val Glu Tyr Ala Glu Leu Ala Ala Glu
    145              150              155              160

Val Ala Gly Val Glu Tyr Val Asp His Trp Ser Tyr Val Asp Ser Ile
    165              170              175

Tyr Glu Thr Leu Gly Asn Ala Thr Val Asn Ser Tyr Phe Pro Ile Asp
    180              185              190

His Thr His Thr Ser Pro Ala Gly Ala Glu Val Val Ala Glu Ala Phe
    195              200              205

Leu Lys Ala Val Val Cys Thr Gly Thr Ser Leu Lys Ser Val Leu Thr
    210              215              220

Thr Thr Ser Phe Glu Gly Thr Cys
    225              230

```

<210> 96

<211> 184

<212> PRT

<213> Escherichia coli

<400> 96

```

Ala Asp Thr Leu Leu Ile Leu Gly Asp Ser Leu Ser Ala Gly Tyr Arg
  1          5          10          15

Met Ser Ala Ser Ala Ala Trp Pro Ala Leu Leu Asn Asp Lys Trp Gln
    20          25          30

Ser Lys Thr Ser Val Val Asn Ala Ser Ile Ser Gly Asp Thr Ser Gln
    35          40          45

Gln Gly Leu Ala Arg Leu Pro Ala Leu Leu Lys Gln His Gln Pro Arg
    50          55          60

Trp Val Leu Val Glu Leu Gly Gly Asn Asp Gly Leu Arg Gly Phe Gln
    65          70          75          80

Pro Gln Gln Thr Glu Gln Thr Leu Arg Gln Ile Leu Gln Asp Val Lys
    85          90          95

```

164

Ala Ala Asn Ala Glu Pro Leu Leu Met Gln Ile Arg Leu Pro Ala Asn
100 105 110

Tyr Gly Arg Arg Tyr Asn Glu Ala Phe Ser Ala Ile Tyr Pro Lys Leu
115 120 125

Ala Lys Glu Phe Asp Val Pro Leu Leu Pro Phe Phe Met Glu Glu Val
130 135 140

Tyr Leu Lys Pro Gln Trp Met Gln Asp Asp Gly Ile His Pro Asn Arg
145 150 155 160

Asp Ala Gln Pro Phe Ile Ala Asp Trp Met Ala Lys Gln Leu Gln Pro
165 170 175

Leu Val Asn His Asp Ser Leu Glu
180

<210> 97

<211> 308

<212> PRT

<213> Aeromonas hydrophila

<400> 97

Ile Val Met Phe Gly Asp Ser Leu Ser Asp Thr Gly Lys Met Tyr Ser
1 5 10 15

Lys Met Arg Gly Tyr Leu Pro Ser Ser Pro Pro Tyr Tyr Glu Gly Arg
20 25 30

Phe Ser Asn Gly Pro Val Trp Leu Glu Gln Leu Thr Asn Glu Phe Pro
35 40 45

Gly Leu Thr Ile Ala Asn Glu Ala Glu Gly Gly Pro Thr Ala Val Ala
50 55 60

Tyr Asn Lys Ile Ser Trp Asn Pro Lys Tyr Gln Val Ile Asn Asn Leu
65 70 75 80

Asp Tyr Glu Val Thr Gln Phe Leu Gln Lys Asp Ser Phe Lys Pro Asp
85 90 95

Asp Leu Val Ile Leu Trp Val Gly Ala Asn Asp Tyr Leu Ala Tyr Gly
100 105 110

Trp Asn Thr Glu Gln Asp Ala Lys Arg Val Arg Asp Ala Ile Ser Asp
115 120 125

Ala Ala Asn Arg Met Val Leu Asn Gly Ala Lys Glu Ile Leu Leu Phe

165

130		135		140
Asn Leu Pro Asp Leu Gly Gln Asn Pro Ser Ala Arg Ser Gln Lys Val				
145		150		155 160
Val Glu Ala Ala Ser His Val Ser Ala Tyr His Asn Gln Leu Leu Leu				
		165		170 175
Asn Leu Ala Arg Gln Leu Ala Pro Thr Gly Met Val Lys Leu Phe Glu				
		180		185 190
Ile Asp Lys Gln Phe Ala Glu Met Leu Arg Asp Pro Gln Asn Phe Gly				
		195		200 205
Leu Ser Asp Gln Arg Asn Ala Cys Tyr Gly Gly Ser Tyr Val Trp Lys				
		210		215 220
Pro Phe Ala Ser Arg Ser Ala Ser Thr Asp Ser Gln Leu Ser Ala Phe				
		225		230 235 240
Asn Pro Gln Glu Arg Leu Ala Ile Ala Gly Asn Pro Leu Leu Ala Gln				
		245		250 255
Ala Val Ala Ser Pro Met Ala Ala Arg Ser Ala Ser Thr Leu Asn Cys				
		260		265 270
Glu Gly Lys Met Phe Trp Asp Gln Val His Pro Thr Thr Val Val His				
		275		280 285
Ala Ala Leu Ser Glu Pro Ala Ala Thr Phe Ile Glu Ser Gln Tyr Glu				
		290		295 300
Phe Leu Ala His				
305				

<210> 98

<211> 167

<212> PRT

<213> Escherichia coli

166

<400> 98

Leu Leu Ile Leu Gly Asp Ser Leu Ser Ala Gly Tyr Arg Met Ser Ala
 1 5 10 15

Ser Ala Ala Trp Pro Ala Leu Leu Asn Asp Lys Trp Gln Ser Lys Thr
 20 25 30

Ser Val Val Asn Ala Ser Ile Ser Gly Asp Thr Ser Gln Gln Gly Leu
 35 40 45

Ala Arg Leu Pro Ala Leu Leu Lys Gln His Gln Pro Arg Trp Val Leu
 50 55 60

Val Glu Leu Gly Gly Asn Asp Gly Leu Arg Gly Phe Gln Pro Gln Gln
 65 70 75 80

Thr Glu Gln Thr Leu Arg Gln Ile Leu Gln Asp Val Lys Ala Ala Asn
 85 90 95

Ala Glu Pro Leu Leu Met Gln Ile Arg Leu Pro Ala Asn Tyr Gly Arg
 100 105 110

Arg Tyr Asn Glu Ala Phe Ser Ala Ile Tyr Pro Lys Leu Ala Lys Glu
 115 120 125

Phe Asp Val Pro Leu Leu Pro Phe Phe Met Glu Glu Val Tyr Leu Lys
 130 135 140

Pro Gln Trp Met Gln Asp Asp Gly Ile His Pro Asn Arg Asp Ala Gln
 145 150 155 160

Pro Phe Ile Ala Asp Trp Met
 165

5

<210> 99

<211> 295

<212> PRT

<213> *Aeromonas hydrophila*

167

<400> 99

Ile Val Met Phe Gly Asp Ser Leu Ser Asp Thr Gly Lys Met Tyr Ser
 1 5 10 15

Lys Met Arg Gly Tyr Leu Pro Ser Ser Pro Pro Tyr Tyr Glu Gly Arg
 20 25 30

Phe Ser Asn Gly Pro Val Trp Leu Glu Gln Leu Thr Asn Glu Phe Pro
 35 40 45

Gly Leu Thr Ile Ala Asn Glu Ala Glu Gly Gly Pro Thr Ala Val Ala
 50 55 60

Tyr Asn Lys Ile Ser Trp Asn Pro Lys Tyr Gln Val Ile Asn Asn Leu
 65 70 75 80

Asp Tyr Glu Val Thr Gln Phe Leu Gln Lys Asp Ser Phe Lys Pro Asp
 85 90 95

Asp Leu Val Ile Leu Trp Val Gly Ala Asn Asp Tyr Leu Ala Tyr Gly

168

100					105					110					
Trp	Asn	Thr	Glu	Gln	Asp	Ala	Lys	Arg	Val	Arg	Asp	Ala	Ile	Ser	Asp
	115						120					125			
Ala	Ala	Asn	Arg	Met	Val	Leu	Asn	Gly	Ala	Lys	Glu	Ile	Leu	Leu	Phe
	130						135				140				
Asn	Leu	Pro	Asp	Leu	Gly	Gln	Asn	Pro	Ser	Ala	Arg	Ser	Gln	Lys	Val
145					150					155					160
Val	Glu	Ala	Ala	Ser	His	Val	Ser	Ala	Tyr	His	Asn	Gln	Leu	Leu	Leu
				165					170					175	
Asn	Leu	Ala	Arg	Gln	Leu	Ala	Pro	Thr	Gly	Met	Val	Lys	Leu	Phe	Glu
			180					185					190		
Ile	Asp	Lys	Gln	Phe	Ala	Glu	Met	Leu	Arg	Asp	Pro	Gln	Asn	Phe	Gly
	195						200					205			
Leu	Ser	Asp	Gln	Arg	Asn	Ala	Cys	Tyr	Gly	Gly	Ser	Tyr	Val	Trp	Lys
210						215					220				
Pro	Phe	Ala	Ser	Arg	Ser	Ala	Ser	Thr	Asp	Ser	Gln	Leu	Ser	Ala	Phe
225					230					235					240
Asn	Pro	Gln	Glu	Arg	Leu	Ala	Ile	Ala	Gly	Asn	Pro	Leu	Leu	Ala	Gln
				245					250					255	
Ala	Val	Ala	Ser	Pro	Met	Ala	Ala	Arg	Ser	Ala	Ser	Thr	Leu	Asn	Cys
			260					265					270		
Glu	Gly	Lys	Met	Phe	Trp	Asp	Gln	Val	His	Pro	Thr	Thr	Val	Val	His
	275						280					285			
Ala	Ala	Leu	Ser	Glu	Pro	Ala									
290						295									

<210> 100

<211> 335

<212> PRT

<213> Aeromonas hydrophila

169

<400> 100

Met	Lys	Lys	Trp	Phe	Val	Cys	Leu	Leu	Gly	Leu	Val	Ala	Leu	Thr	Val
1				5					10					15	

Gln	Ala	Ala	Asp	Ser	Arg	Pro	Ala	Phe	Ser	Arg	Ile	Val	Met	Phe	Gly
			20					25					30		

170

Asp Ser Leu Ser Asp Thr Gly Lys Met Tyr Ser Lys Met Arg Gly Tyr
 35 40 45
 Leu Pro Ser Ser Pro Pro Tyr Tyr Glu Gly Arg Phe Ser Asn Gly Pro
 50 55 60
 Val Trp Leu Glu Gln Leu Thr Asn Glu Phe Pro Gly Leu Thr Ile Ala
 65 70 75 80
 Asn Glu Ala Glu Gly Gly Pro Thr Ala Val Ala Tyr Asn Lys Ile Ser
 85 90 95
 Trp Asn Pro Lys Tyr Gln Val Ile Asn Asn Leu Asp Tyr Glu Val Thr
 100 105 110
 Gln Phe Leu Gln Lys Asp Ser Phe Lys Pro Asp Asp Leu Val Ile Leu
 115 120 125
 Trp Val Gly Ala Asn Asp Tyr Leu Ala Tyr Gly Trp Asn Thr Glu Gln
 130 135 140
 Asp Ala Lys Arg Val Arg Asp Ala Ile Ser Asp Ala Ala Asn Arg Met
 145 150 155 160
 Val Leu Asn Gly Ala Lys Glu Ile Leu Leu Phe Asn Leu Pro Asp Leu
 165 170 175
 Gly Gln Asn Pro Ser Ala Arg Ser Gln Lys Val Val Glu Ala Ala Ser
 180 185 190
 His Val Ser Ala Tyr His Asn Gln Leu Leu Leu Asn Leu Ala Arg Gln
 195 200 205
 Leu Ala Pro Thr Gly Met Val Lys Leu Phe Glu Ile Asp Lys Gln Phe
 210 215 220
 Ala Glu Met Leu Arg Asp Pro Gln Asn Phe Gly Leu Ser Asp Gln Arg
 225 230 235 240
 Asn Ala Cys Tyr Gly Gly Ser Tyr Val Trp Lys Pro Phe Ala Ser Arg
 245 250 255
 Ser Ala Ser Thr Asp Ser Gln Leu Ser Ala Phe Asn Pro Gln Glu Arg
 260 265 270
 Leu Ala Ile Ala Gly Asn Pro Leu Leu Ala Gln Ala Val Ala Ser Pro
 275 280 285

171

Met Ala Ala Arg Ser Ala Ser Thr Leu Asn Cys Glu Gly Lys Met Phe
290 295 300

Trp Asp Gln Val His Pro Thr Thr Val Val His Ala Ala Leu Ser Glu
305 310 315 320

Pro Ala Ala Thr Phe Ile Glu Ser Gln Tyr Glu Phe Leu Ala His
325 330 335

<210> 101

<211> 318

<212> PRT

5 <213> *Aeromonas salmonicida*

<400> 101

Ala Asp Thr Arg Pro Ala Phe Ser Arg Ile Val Met Phe Gly Asp Ser
1 5 10 15

Leu Ser Asp Thr Gly Lys Met Tyr Ser Lys Met Arg Gly Tyr Leu Pro
20 25 30

Ser Ser Pro Pro Tyr Tyr Glu Gly Arg Phe Ser Asn Gly Pro Val Trp
35 40 45

Leu Glu Gln Leu Thr Lys Gln Phe Pro Gly Leu Thr Ile Ala Asn Glu
50 55 60

Ala Glu Gly Gly Ala Thr Ala Val Ala Tyr Asn Lys Ile Ser Trp Asn
65 70 75 80

Pro Lys Tyr Gln Val Tyr Asn Asn Leu Asp Tyr Glu Val Thr Gln Phe
85 90 95

Leu Gln Lys Asp Ser Phe Lys Pro Asp Asp Leu Val Ile Leu Trp Val
100 105 110

Gly Ala Asn Asp Tyr Leu Ala Tyr Gly Trp Asn Thr Glu Gln Asp Ala
115 120 125

Lys Arg Val Arg Asp Ala Ile Ser Asp Ala Ala Asn Arg Met Val Leu
130 135 140

Asn Gly Ala Lys Gln Ile Leu Leu Phe Asn Leu Pro Asp Leu Gly Gln
145 150 155 160

Asn Pro Ser Ala Arg Ser Gln Lys Val Val Glu Ala Val Ser His Val
165 170 175

Ser Ala Tyr His Asn Lys Leu Leu Leu Asn Leu Ala Arg Gln Leu Ala
180 185 190

172

Pro Thr Gly Met Val Lys Leu Phe Glu Ile Asp Lys Gln Phe Ala Glu
195 200 205

Met Leu Arg Asp Pro Gln Asn Phe Gly Leu Ser Asp Val Glu Asn Pro
210 215 220

Cys Tyr Asp Gly Gly Tyr Val Trp Lys Pro Phe Ala Thr Arg Ser Val
225 230 235 240

Ser Thr Asp Arg Gln Leu Ser Ala Phe Ser Pro Gln Glu Arg Leu Ala
245 250 255

Ile Ala Gly Asn Pro Leu Leu Ala Gln Ala Val Ala Ser Pro Met Ala
260 265 270

Arg Arg Ser Ala Ser Pro Leu Asn Cys Glu Gly Lys Met Phe Trp Asp
275 280 285

Gln Val His Pro Thr Thr Val Val His Ala Ala Leu Ser Glu Arg Ala
290 295 300

Ala Thr Phe Ile Glu Thr Gln Tyr Glu Phe Leu Ala His Gly
305 310 315

<210> 102

<211> 50

5 <212> PRT

<213> Artificial

<220>

<223> Consensus sequence

<400> 102

Arg Pro Ala Phe Ser Arg Ile Val Met Phe Gly Asp Ser Leu Ser Asp
1 5 10 15

Thr Gly Lys Met Tyr Ser Lys Met Arg Gly Tyr Leu Pro Ser Ser Pro
20 25 30

Pro Tyr Tyr Glu Gly Arg Phe Ser Asn Gly Pro Val Trp Leu Glu Gln
35 40 45

Leu Thr
50

10

173

<210> 103

< 211> 13

< 212> PRT

< 213> Artificial

5

<220>

< 223> Consensus sequence

<400> 103

Phe	Pro	Gly	Leu	Thr	Ile	Ala	Asn	Glu	Ala	Glu	Gly	Gly
1				5					10			

<210> 104

10

< 211> 79

< 212> PRT

< 213> Artificial

<220>

< 223> Consensus sequence

15

<400> 104

Thr	Ala	Val	Ala	Tyr	Asn	Lys	Ile	Ser	Trp	Asn	Pro	Lys	Tyr	Gln	Val
1				5					10					15	

Ile	Asn	Asn	Leu	Asp	Tyr	Glu	Val	Thr	Gln	Phe	Leu	Gln	Lys	Asp	Ser
			20					25					30		

Phe	Lys	Pro	Asp	Asp	Leu	Val	Ile	Leu	Trp	Val	Gly	Ala	Asn	Asp	Tyr
		35					40					45			

Leu	Ala	Tyr	Gly	Trp	Asn	Thr	Glu	Gln	Asp	Ala	Lys	Arg	Val	Arg	Asp
	50					55					60				

Ala	Ile	Ser	Asp	Ala	Ala	Asn	Arg	Met	Val	Leu	Asn	Gly	Ala	Lys
65						70					75			

<210> 105

< 211> 23

< 212> PRT

20

< 213> Artificial

174

<220>

< 223> Consensus sequence

<400> 105

Ile Leu Leu Phe Asn Leu Pro Asp Leu Gly Gln Asn Pro Ser Ala Arg
 1 5 10 15

Ser Gln Lys Val Val Glu Ala
 20

5 <210> 106

< 211> 8

< 212> PRT

< 213> Artificial

<220>

10 < 223> Consensus sequence

<400> 106

Ser His Val Ser Ala Tyr His Asn
 1 5

<210> 107

< 211> 38

15 < 212> PRT

< 213> Artificial

<220>

< 223> Consensus sequence

<400> 107

Leu Leu Leu Asn Leu Ala Arg Gln Leu Ala Pro Thr Gly Met Val Lys
 1 5 10 15

Leu Phe Glu Ile Asp Lys Gln Phe Ala Glu Met Leu Arg Asp Pro Gln
 20 25 30

20 Asn Phe Gly Leu Ser Asp
35

<210> 108

< 211> 7

175

< 212> PRT

< 213> Artificial

<220>

< 223> Consensus sequence

5 <400> 108

Tyr	Val	Trp	Lys	Pro	Phe	Ala
1				5		

<210> 109

< 211> 5

< 212> PRT

10 < 213> Artificial

<220>

< 223> Consensus sequence

<400> 109

Gln	Leu	Ser	Ala	Phe
1			5	

15 <210> 110

< 211> 22

< 212> PRT

< 213> Artificial

<220>

20 < 223> Consensus sequence

<400> 110

Pro	Gln	Glu	Arg	Leu	Ala	Ile	Ala	Gly	Asn	Pro	Leu	Leu	Ala	Gln	Ala
1				5				10					15		

Val	Ala	Ser	Pro	Met	Ala
			20		

<210> 111

< 211> 4

176

< 212> PRT

< 213> Artificial

<220>

< 223> Consensus sequence

5 <400> 111

Arg Ser Ala Ser
1

<210> 112

< 211> 24

< 212> PRT

10 < 213> Artificial

<220>

< 223> Consensus sequence

<400> 112

Leu Asn Cys Glu Gly Lys Met Phe Trp Asp Gln Val His Pro Thr Thr
1 5 10 15

Val Val His Ala Ala Leu Ser Glu
20

15 <210> 113

< 211> 5

< 212> PRT

< 213> Artificial

<220>

20 < 223> Consensus sequence

<400> 113

Ala Ala Thr Phe Ile
1 5

<210> 114

< 211> 7

< 212> PRT

< 213> Artificial

 $\langle 220 \rangle$

< 223> Consensus sequence

5 <400> 114

Gln Tyr Glu Phe Leu Ala His
1 5

<210> 115

 $\langle 211 \rangle$ 1225

< 212> DNA

10 < 213> Artificial

 $\langle 220 \rangle$

< 223> XhoI insert containing the LAT-KLM3' precursor gene

 $\langle 220 \rangle$

< 221> CDS

15 $\langle 222 \rangle$ (101)..(1144)

<400> 115

```

gctttttcttt tggaagaaaa tatagggaaa atggtacttg ttaaaaattc ggaatattta      60
tacaatatca tatgtttcac attgaaaggg gaggagaatc atg aaa caa caa aaa      115
                                         Met Lys Gln Gln Lys
                                         1           5

cgg ctt tac gcc cga ttg ctg acg ctg tta ttt gcg ctc atc ttc ttg      163
Arg Leu Tyr Ala Arg Leu Leu Thr Leu Leu Phe Ala Leu Ile Phe Leu
                        10                15                20

ctg cct cat tct gca gct tca gca gca gat aca aga ccg gcg ttt agc      211
Leu Pro His Ser Ala Ala Ser Ala Ala Asp Thr Arg Pro Ala Phe Ser
                        25                30                35

cgg atc gtc atg ttt gga gat agc ctg agc gat acg gcc aaa atg tat      259
Arg Ile Val Met Phe Gly Asp Ser Leu Ser Asp Thr Gly Lys Met Tyr
                        40                45                50

agc aaa atg aga ggc tat ctt ccg tca agc ccg ccg tat tat gaa ggc      307
Ser Lys Met Arg Gly Tyr Leu Pro Ser Ser Pro Pro Tyr Tyr Glu Gly
                        55                60                65

cgc ttt agc aat gga ccg gtc tgg ctg gaa caa ctg acg aaa caa ttt      355
Arg Phe Ser Asn Gly Pro Val Trp Leu Glu Gln Leu Thr Lys Gln Phe
                        70                75                80                85

ccg gga ctg acg atc gct aat gaa gca gaa gga gga gca aca gcg gtc      403
Pro Gly Leu Thr Ile Ala Asn Glu Ala Glu Gly Gly Ala Thr Ala Val
                        90                95                100

gcc tat aac aaa atc agc tgg gac ccg aaa tat cag gtc atc aac aac      451
Ala Tyr Asn Lys Ile Ser Trp Asp Pro Lys Tyr Gln Val Ile Asn Asn
                        105                110                115

ctg gac tat gaa gtc aca cag ttt ctt cag aaa gac agc ttt aaa ccg      499
Leu Asp Tyr Glu Val Thr Gln Phe Leu Gln Lys Asp Ser Phe Lys Pro
                        120                125                130

gat gat ctg gtc atc ctt tgg gtc gcc gcc aat gat tat ctg gcg tat      547
Asp Asp Leu Val Ile Leu Trp Val Gly Ala Asn Asp Tyr Leu Ala Tyr
                        135                140                145

ggc tgg aac aca gaa caa gat gcc aaa aga gtc aga gat gcc atc agc      595
Gly Trp Asn Thr Glu Gln Asp Ala Lys Arg Val Arg Asp Ala Ile Ser
                        150                155                160                165

gat gcc gct aat aga atg gtc ctg aac gcc gcc aaa caa atc ctg ctg      643

```

179

Asp	Ala	Ala	Asn	Arg	Met	Val	Leu	Asn	Gly	Ala	Lys	Gln	Ile	Leu	Leu	
				170					175					180		
ttt	aac	ctg	ccg	gat	ctg	gga	caa	aat	ccg	agc	gcc	aga	agc	caa	aaa	691
Phe	Asn	Leu	Pro	Asp	Leu	Gly	Gln	Asn	Pro	Ser	Ala	Arg	Ser	Gln	Lys	
			185					190					195			
gtc	gtc	gaa	gca	gtc	agc	cat	gtc	agc	gcc	tat	cat	aac	aaa	ctg	ctg	739
Val	Val	Glu	Ala	Val	Ser	His	Val	Ser	Ala	Tyr	His	Asn	Lys	Leu	Leu	
		200					205					210				
ctg	aac	ctg	gca	aga	caa	ttg	gca	ccg	acg	gga	atg	gtt	aaa	ttg	ttt	787
Leu	Asn	Leu	Ala	Arg	Gln	Phe	Ala	Pro	Thr	Gly	Met	Val	Lys	Leu	Phe	
	215					220					225					
gaa	att	gac	aaa	cag	ttt	gcc	gaa	atg	ctg	aga	gat	ccg	caa	aat	ttt	835
Glu	Ile	Asp	Lys	Gln	Phe	Ala	Glu	Met	Leu	Arg	Asp	Pro	Gln	Asn	Phe	
230				235					240					245		
ggc	ctg	agc	gat	gtc	gaa	aac	ccg	tgc	tat	gat	ggc	gga	tat	gtc	tg	883
Gly	Leu	Ser	Asp	Val	Glu	Asn	Pro	Cys	Tyr	Asp	Gly	Gly	Tyr	Val	Trp	
			250					255						260		
aaa	ccg	ttt	gcc	aca	aga	agc	gtc	agc	acg	gat	aga	caa	ctg	tca	gcg	931
Lys	Pro	Phe	Ala	Thr	Arg	Ser	Val	Ser	Thr	Asp	Arg	Gln	Leu	Ser	Ala	
			265					270					275			
ttt	agc	ccg	caa	gaa	aga	ctg	gca	atc	gcc	gga	aat	ccg	ctt	ttg	gca	979
Phe	Ser	Pro	Gln	Glu	Arg	Leu	Ala	Ile	Ala	Gly	Asn	Pro	Leu	Leu	Ala	
		280				285						290				
caa	gca	gtt	gct	tca	ccg	atg	gca	aga	aga	tca	gca	agc	ccg	ctg	aat	1027
Gln	Ala	Val	Ala	Ser	Pro	Met	Ala	Arg	Arg	Ser	Ala	Ser	Pro	Leu	Asn	
	295					300					305					
tgc	gaa	ggc	aaa	atg	ttt	tg	gat	cag	gtc	cat	ccg	aca	aca	gtt	gtc	1075
Cys	Glu	Gly	Lys	Met	Phe	Trp	Asp	Gln	Val	His	Pro	Thr	Thr	Val	Val	
310				315					320					325		
cat	gct	gcc	ctt	tca	gaa	aga	gcg	gcg	acg	ttt	atc	gaa	aca	cag	tat	1123
His	Ala	Ala	Leu	Ser	Glu	Arg	Ala	Ala	Thr	Phe	Ile	Glu	Thr	Gln	Tyr	
			330					335						340		
gaa	ttt	ctg	gcc	cat	ggc	tga	gttaacagag	gacggatttc	ctgaaggaaa							1174
Glu	Phe	Leu	Ala	His	Gly											
			345													
tccggtttttt	tatttttaagc	ttggagacaa	ggtaaaggat	aaaacctcga	g											1225

<210> 116

<211> 347

<212> PRT

<213> Artificial

<220>

<223> Synthetic Construct

180

<400> 116

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 Ser Ala Ala Asp Thr

20 25 30

Arg Pro Ala Phe Ser Arg Ile Val Met Phe Gly Asp Ser Leu Ser Asp
 35 40 45

Thr Gly Lys Met Tyr Ser Lys Met Arg Gly Tyr Leu Pro Ser Ser Pro
 50 55 60

Pro Tyr Tyr Glu Gly Arg Phe Ser Asn Gly Pro Val Trp Leu Glu Gln
 65 70 75 80

Leu Thr Lys Gln Phe Pro Gly Leu Thr Ile Ala Asn Glu Ala Glu Gly
 85 90 95

Gly Ala Thr Ala Val Ala Tyr Asn Lys Ile Ser Trp Asp Pro Lys Tyr
 100 105 110

Gln Val Ile Asn Asn Leu Asp Tyr Glu Val Thr Gln Phe Leu Gln Lys
 115 120 125

Asp Ser Phe Lys Pro Asp Asp Leu Val Ile Leu Trp Val Gly Ala Asn
 130 135 140

Asp Tyr Leu Ala Tyr Gly Trp Asn Thr Glu Gln Asp Ala Lys Arg Val
 145 150 155 160

Arg Asp Ala Ile Ser Asp Ala Ala Asn Arg Met Val Leu Asn Gly Ala
 165 170 175

Lys Gln Ile Leu Leu Phe Asn Leu Pro Asp Leu Gly Gln Asn Pro Ser
 180 185 190

Ala Arg Ser Gln Lys Val Val Glu Ala Val Ser His Val Ser Ala Tyr
 195 200 205

His Asn Lys Leu Leu Leu Asn Leu Ala Arg Gln Leu Ala Pro Thr Gly
 210 215 220

Met Val Lys Leu Phe Glu Ile Asp Lys Gln Phe Ala Glu Met Leu Arg
 225 230 235 240

Asp Pro Gln Asn Phe Gly Leu Ser Asp Val Glu Asn Pro Cys Tyr Asp
 245 250 255

Gly Gly Tyr Val Trp Lys Pro Phe Ala Thr Arg Ser Val Ser Thr Asp
 260 265 270

Arg Gln Leu Ser Ala Phe Ser Pro Gln Glu Arg Leu Ala Ile Ala Gly
 275 280 285

181

Asn Pro Leu Leu Ala Gln Ala Val Ala Ser Pro Met Ala Arg Arg Ser
 290 295 300

Ala Ser Pro Leu Asn Cys Glu Gly Lys Met Phe Trp Asp Gln Val His
 305 310 315 320

Pro Thr Thr Val Val His Ala Ala Leu Ser Glu Arg Ala Ala Thr Phe
 325 330 335

Ile Glu Thr Gln Tyr Glu Phe Leu Ala His Gly
 340 345

<210> 117

<211> 4

5 <212> PRT

<213> Artificial

<220>

<223> Sequence motif

<220>

10 <221> SITE

<222> (4) .. (4)

<223> Xaa may be one or more of the following amino acid residues L, A, V, I, F, Y, H, Q, T, N, M
 or S.

<400> 117

15 Gly Asp Ser Xaa
1

<210> 118

<211> 5

<212> PRT

<213> Artificial

20 <220>

<223> Sequence motif

<220>

<221> SITE

182

< 222> (2) .. (2)
< 223> May be Gly

<220>
< 221> SITE
< 222> (5)..(5)
< 223> May be Ala or Leu

<400> 118
Gly Ala Asn Asp Tyr
1 5

5

10

Patentkrav

1. Fremgangsmåde til fremstilling af en lipid acyltransferase, hvilken fremgangsmåde omfatter trinnene:
 - 5 (i) tilvejebringelse af en *Bacillus*-celle, hvor *Bacillus*-cellen er *B. licheniformis*;
 - (ii) omdannelse af *Bacillus*-cellen, hvor *Bacillus*-cellen er *B. licheniformis*, med den heterolog nukleotidsekvens, der koder for en lipid acyltransferase ifølge SEQ ID NO: 49; og
 - (iii) ekspresion af den lipide acyltransferase i cellen under styring af en promotersekvens.
- 10 2. Fremgangsmåde ifølge krav 1, hvor promotersekvensen ikke er nativt forbundet med den nukleotidsekvens, der koder for en lipid acyltransferase.
3. Fremgangsmåde ifølge et hvilket som helst af de foregående krav, hvor fremgangsmåden omfatter det supplerende trin med isolering/genvinding af den lipide acyltransferase.
- 15 4. Fremgangsmåde ifølge et hvilket som helst af de foregående krav, hvor promotersekvensen er homolog med værtscellen.
5. Fremgangsmåde ifølge et hvilket som helst af de foregående krav, hvor promotersekvensen er udvalgt fra gruppen bestående af en α -amylase-promotersekvens, en protease-promotersekvens, en subtilisin-promotersekvens, en glutaminsyrespecifik protease-promotersekvens og en levansucrase-promotersekvens.
- 20 6. Lipid acyltransferase, der kan opnås ved hjælp af fremgangsmåden ifølge et hvilket som helst af de foregående krav, og som kodes for af SEQ ID NO: 49.
- 25

Drawing

FIGURE 1 (SEQ ID No. 16)

```

1  ADTREAFAFSRI VMFGDSLSDT GKMYSKMRGY LPSSPPYYEG RFSNGCPVWLE QLTQKQFPLT
61  IANEAEAGGAT AVAYNKISWD PKYQVINNLD YEVTOFLQKD SFKPDDELVL WVGVANDYLAY
121 GWNTEQDAKR VRDAISDAAN RMVLNGAKQI LLFNLPDLGQ NPSARSQKV VAVSHVSAYH
181 NKLLNLARQ LAPTMGVKLF EIDKQFAEML RDPQNFGLSD VENPCYDGGY VWKPFATRSV
241 STDRQLSAFS PQERLAIAGN PLLAQAVASP MARRSASPLN CEGRMFWDQV HPTTVVHAAL
301 SERAATFIET QYEFLAHG

```

FIGURE 2 (SEQ ID No. 1)

```

1  MKKWFVCLLG LVALTVQAAD SRPAFSRIVM FGDSLSDTGK MYSKMRGYLP
51  SSPPYYEGRF SNGPVWLEQL TKQFPGLTIA NEAEGGATAV AYNKISWNPK
101 YQVINNLDYE VTQFLQKDSF KPDDLVLWV GANDYLAYGW NTEQDAKRVR
151 DAISDAANRM VINGAKQILL FNLPDLGQNP SRSQKVVEA VSHVSAYHNQ
201 LLLNLARQLA PTGMVKLFEL DKQFAEMLRD PQNFGLSDE NPCYDGGYVW
251 KPFATRSVST DRQLSAFSPQ ERLAIAGNPL LAQAVASPM RRSASPLNCE
301 GKMFWDDQVHP TTVVHAALSE RAATFIANQY EFLAH*

```

5

FIGURE 3 (SEQ ID No. 2)

```

1  ivafGD51LTd geayygsdsd ggwgagiadr Ltallrlrar prgydvfnrg isGrtsdGrl
61  ivDalvallf laqslglpnL pPYLsgdflr GANFAsagAt IlptsgpflI QvqFkdifksc
121 vlelrqalgI lqallrlpv ldkaspdlvt imiGtNDlit saffgpkste sdrnsvvpef
181 kdnlrqlikr lrsnngarii vlitlvilnl gplGClPlkl alalassknv dasgclerln
241 eavadfneal relaiskied qlrkdgldpv kgadvpyvDl ysifgdldgi gnpsayvyGF
301 ettkaCCGyG gryNynzvCG naglcnvtaK aCnpssylls flfwDgiHps ekGykavAea
361 1

```

FIGURE 4 (SEQ ID No. 3)

```

1  mkkwfvcllg lvaltvqaad srpafsrivm fgdsldstgk myskmrgylp ssppyyegrf
61  sngpvwleql tnefpgltia neaeggptav aynkiswnpk yqvinnldye vtqflqkdsf
121 kpddlvlwv gandylaygw nteqdakrvr daisdaanrm vlngakeill fnlpdlgqnp
181 sarsqkvvea ashvsayhnq lllnlarqla ptgmvkfeli dkqfaemlrd pqnfglsdqr
241 nacyggsyvw kpfasrsast dsqlsafnpq erlaiagnpl laqavaspma arsastlnce
301 gkmfwdqvhv ttvvhaalse paatfiesgy eflah

```

FIGURE 5 (SEQ ID No. 4)

```

1 mkkwfvcllg lialtvqaad trpaferiwm fgdsiesdtgk myskmrqyyp sappyyegxf
61 sngpvwleql tkqfpgltia neaeggataav aynkiswnpk yqvynnldye vtqflqkdsf
121 kpddlvilwv gandylyaygw ntegdakrvr daisdeanm vlngakqill fnlpdlggnp
181 sarsqkvvea vshvsayhmk llnlarqla ptgmwklfei dkqfaemlrp pqnfcldve
241 npcydgggyvw kpfatrsvst drqlsafspq azlaiagnpl laqavaspmr rrsasplnce
301 gbmfwdqvhv ttvvhaaalse raatfietcy eflahg

```

FIGURE 6 (SEQ ID No. 5)

```

1 mpkpalrrvm tatvaavgtl algltdatab aapaqatptl dyvalgdsys agsgvlpvdp
61 anlcllrsta nyphviadt garltdvtcg aaqtadftta qypgvapqld algtgtdlvt
121 ltiggnndst finaitacgt agvlsqgkgs pckdrhgtsf ddeieantyp alkeallgvr
181 arapharvaa lgyppwitpat adpscfllkp laagdvpplr aqahlnlav rraaetgat
241 yvdfsgvsdg hdaceapgtr wiepllfghs lvpvhpnaig ermaehtmd vlgld

```

FIGURE 7 (SEQ ID No. 6)

```

1 mpkpalrrvm tatvaavgtl algltdatab aapaqatptl dyvalgdsys agsgvlpvdp
61 anlcllrsta nyphviadt garltdvtcg aaqtadftta qypgvapqld algtgtdlvt
121 ltiggnndst finaitacgt agvlsqgkgs pckdrhgtsf ddeieantyp alkeallgvr
181 arapharvaa lgyppwitpat adpscfllkp laagdvpplr aqahlnlav rraaetgat
241 yvdfsgvsdg hdaceapgtr wiepllfghs lvpvhpnaig ermaehtmd vlgld

```

FIGURE 8 (SEQ ID No. 7)

```

1 mdyekfllfg dsitefafnt rpiedgkdgy algaalvney trkmdilqrg fxgytsrwai
61 kilpeilkhe snivmatifl gandacsagp qsvplpefid nirqmvslnk syhixpiig
121 pglvdrekwe kekseeialg yftrnenfai ysdalaklan eekvpfvaln kafqgegga
181 wqqltdqih fsgkgykifh dellkviets ypyhpknmq yklkdwrclv ddgsnms

```

FIGURE 9 (SEQ ID No. 8)

10	20	30	40	50	60
MNLRQWGA	TAALALGLAA	CGGGGTDQSG	NPNVAKVQEM	VWFGDSLSDI	GTYPVPAQAV
70	80	90	100	110	120
GGGKFTTNPG	PIKAEIVAAQ	LGVTLPVAVM	GYATSVQNC	KAGCFDYAQG	GSRVTDPMGI
130	140	150	160	170	180
GHNGCAGALT	YFVQQQLANF	YLAASNTFNG	NNDVVEVLAG	SNDIFFWTTA	AATSGSGVTP
190	200	210	220	230	240
AIATAQVQQA	ANDLVGYVKD	MIKAGATQVY	VFNLPDSSLT	PDGVASGTTG	QALLHALVET
250	260	270	280	290	300
FWTTLQSGLA	GTSARIIDFN	AQLTAAIQNG	ASFGFANTSA	KACDATKINA	LVPSAGGSSL
310	320	330	340		
FCSANTLVAS	GADQSYLFAD	GVHPTTAGHR	LIASNVLAEL	IADNVAN	

FIGURE 10 (SEQ ID No. 9)

```

1 migsyvavgd sftegvvgdpg pdgafvvgwad rlawlladrr pegdftytnl avrgllldqi
61 vaeqvprvvg lapdlvsfaa ggndiirpgt dpdevaerfe lavaaltaaa gtvltvtgfd
121 trgvvpvlkhl rgklatyngh vraiadrygc pvidlwslls vqdrxawdad rhlhspeght
181 rvalraggal girvpadpdg pwpplpprgt ldvrrddvhw areylvpwig rrlrgessgd
241 hvtakgtlsp daiktriaav a

```

FIGURE 11 (SEQ ID No. 10)

```

1 mqtncpytsl vavgdsfteg msdlldpgsy rgwadllatr maarspggfry anlavrgkli
61 ggiydeqvdr aaamgadvit lvgglndtlr pkcdmarvrd lltqaverla phceqlvlmr
121 spgrqgqvle rfrpmealf aviddlagrh gavvvdlyga gsladprmwv vdrihltaeg
181 hrrvaeavwg slghepedpe whapipatpp pgwvtrrtad vrfarqhlip wigrrltgrs
241 sgdlpakrp dlipyedpar

```

FIGURE 12 (SEQ ID No. 11)

```

1 mtrgrdggag apptkhrall aaivtlivai saaiyagasa ddgsrdhalq aggrlprgda
61 apastgawvg awatapaaae pgtettglag rsvrvvhts vgggtgaritl snlyggsplt
121 vthasialaa gpdtaaaiaa tmrrltfggs arviipaggg vmsdtarlai pyganvlttt
181 yspipsgpvt yhpqarqtsy ladgdrtadv tavayttptp ywryltaldv lsheadgtvv
241 afgdsitdga rsgsdanhrw tdvlaarlhe aagdgdrtpy ysvvnegisg nrlitsrpgp
301 padnpsglsr fqrdrvlerth vkavvvvlgv ndvlinspela drdailtqlr tlvdraharg
361 lrvvgatitp fgggygytea retmrqevne eirsgrvidt vvcfdkalrd pydprmrarsd
421 ydsgdhlhpg dkgyamgav idlaalkgae pvka

```

FIGURE 13 (SEQ ID No. 12)

```

1 mtsmsrarva rriaagaayg gggiglagaa avglvvaevq larrxrvvgvt ptrvpnaqgl
61 yggtlptagd pplrlmalgd staaggqvhr agqtpgalla sglaavaerp vrlgsvaqpg
121 acsddldrqv alvlaepdrv pdicvimvga ndvthmpat rsvrhlssav rrlrtagaev
181 vvgtcpdlgt iervrgplr w larrasrqla aagtigaveq ggrrvslgdl lgpefaqnpr
241 elfgpcnyhp saegyataam avlpsvcaal glwpadeehp dalrragflp varaaacaas
301 eaqtevaam ptgprgwal lkrirrrrrvs eaepspsgy

```

FIGURE 14 (SEQ ID No. 13)

```

1 mgrgtgqrtr ygrrrarval aaltaavlgv gvagcdsvgg dspapsqgsp krtxtapawd
61 tspasvaavg dsitrgfdac avlsdcpevs watgssakvd slavrllgka daaehswnya
121 vtgarmadlt agvtraaqre pelvavmaga ndacrsttsa mtpvadifraq feeamatlrk
181 klpkagvyvs sipdkrlws qgrtnplgkq vwklglcpm lgdadsldse atlrntvrd
241 rvadynevir evcakdrccr sddgavhefr fgtddqlshwd wfhpssvdgga rlaelayrav
301 taknp

```

188

FIGURE 15 (SEQ ID No. 14)

```

1 mrlsrraata sallltpala lfgasaavsa priqatdyva lgdsyssgvg agsydsssgs
61 ckrstkstypa lwaashigtr fnftacsgar tgdvlakgit pvnsgtdlvs itiggnadagf
121 admtttnlq gesaclaria karayiqqtl paqldqvyda idstapaaqv vvlgyprfyk
181 lggscavqls eksraaina addinavtak raadhgfafg dvnttfaghe lcsgapwlhs
241 vtlpvensyh ptanggskgy lpvlnsat

```

FIGURE 16 (SEQ ID No. 15)

```

1 MKKWFVCLLG LIALTVQAAD TRPAFSRIVM FGDSLSDTGK MYSKMRGYLP
51 SSPPYYEGRF SNGPVWLEQL TKQFPGLTIA NEAEGGATAV AYNKISWNP
101 YQVINNL DYE VTQFLQKDSF KPDDLVLWV GANDYLAYGW NTEQDAKRVR
151 DATSDAANRM VLNQAKQIL L FNLPDLGQNP SRSQKVVEA VSHVSAYHNK
201 LLLNLARQLA PTGMVKLFEL DKQFAEMLRD PQNFGLSDVE NPCYDGGYVW
251 KPFATRSVST DRQLSAFSPQ ERLAIAGNPL LAQAVASPMR RRSASPLNCE
301 GKMFWDQVHP TTVVEAALSE RAATFIETQY EFLANG*

```

FIGURE 17 (SEQ ID No. 19)

```

1 migsyvavgd sftegvgdpg pdgafvgwad rlvllldrr pegdftytnl avrgrlldqi
61 vaegvprvvg lapdlvsfaa ggndiirpgt dpdevaerfe lavaaltaaa gtvlvttgfd
121 trgvplkhl rgkistyngh vraiadrygc pvidlwslrs vqdrrawd ad rhlspgght
181 rvalraggal glrvpadpdc pwpplprgt ldvrddvhw areylvpwig nrlrgessgd
241 hvtaqgtlsp daiktriaav a

```

FIGURE 18 (SEQ ID No. 25)

```

1 MFKFKKNFLV GLSRAALMSIS LFSATASAAS ADSREAFSRI VMFGDSLSDT
51 GKMYSKMRGY LPSSPPYYEG RFSNGPVWLE QLTKQFPGLT IANEAEGGAT
101 AVAYNKISWN PKYQVINNL D YEVTOFLQKD SFKPDDLVLV WVGANDYLAY
151 GWNTEQDAKR VRDAISDAAN RMVLNGAKOI LLENLPDLGQ NPSARSQVW
201 EAVSHVSAYH NQLLLMLARQ LAPTMVKLF EIDKQFAEML RDPQNFGLSD
251 VENPCYDGGY VWKPFATRSV STDRQLSAFS PQERLAIAGN PLLAQAVASP
301 MARRSASPLN CEGKMFWDQV HPTTVVHAAL SERAATTIAN QYEFLAH**

```

FIGURE 19 (SEQ ID NO. 26)

```

MRLTRLSAASFVVFALLLALLGISPAQAACPAYVALGDSYSSCNGAGSYIDS3SGDCHRSN
NAYPARWAAANAPSSFTFAACSGAVTTDVINNQLGALNASTGLVSIITIGNDAGFADAMTT
CVTSSDSTCINRLATATNYINTTLLARLDVYSQIKARAPNARVVVLGYPRMYLASNPWYC
LGLSNTKRAATNTTADTLNSVSSRATAHGFRFGDVRPTFNNHELFEGNDWLHSLTLPVWE
SYHFTSTGHQSGYLEPVLNANSST

```

FIGURE 20 (SEQ ID No. 27)

ZP 00058717

```

1 mlphpagerg evgaaffallv gtpqdrirl ecbetrplrg rcgcgerirvp pltlpgdgvl
61 cttsstrdae tvwrkhlqpr pdggfrrphlg vgcilagggs pgvllwcgrrg crfevcrddt
121 pglstrngd ssppfragws lppkcgelsq sarktpavpr yslrtdrpd gprgrfvgsq
181 praatrrrlf lgipalvlvt altlvlavpt gretlwrnwc eatgdwclgv pvdsergqpa
241 dgefillspv qaatwgnyya lgdsyssgdg ardyypgtav kggcwsana ypelvaeayd
301 faghlsflac sggrgyemld aidevgsqld wnsphstlvt igiggndlgf stvlktcmvr
361 vplldskact dgedairkzm akfettfeel isevrtrapd arilvvgypz ifpeeptgay
421 ylttasnqrw lnetigefnq qlaeavavhd eeieasggvg svefvdvyha ldgheigsde
481 pwngvqlrd latgvtvdrs tfnpnaaghr avgarvlegi etgpgxplya tiavvagatv
541 dtlagevg

```

FIGURE 21 (SEQ ID No. 28)

```

1 mgsqpraatr rrlflgipal vlvtaltlvt avptgretlw xmwceatqdw clgvvpdsrg
61 gpaedgefl1 lspvqaetwg nyyalgdsys sgdgardyyt gtavkggcwr sanaypelva
121 eaydfaghls flacsggrgy amldaidevg sqldwnspht slvtigiggn dlqfstvlkt
181 cmvrvp1lds kactdyedai rkrmakfett feelisevrt rapdarilvv gypxifpeep
241 tgayytltas ngrwlnetig efnqqlaeav avhdeeiass ggvgsvfvd vyhaldghei
301 gsdepwvngv qlrdlatgvt vdrstfhpna aghravgerv leqietgpgr plyatfavva
361 gatvdtlage vg

```

FIGURE 22 (SEQ ID No. 29)

```

1 mrttviaasa lillagcadg areetagapp gessggiree gaeastsitd vyialgdsya
61 amgrgdqplr gepfclrssg nypelihaev tdltcggavt gdlleprrtlg ertlpaqvda
121 ltedttlvtl siggndlgfg evagcrreri agenaddcvd lletigeql dqlppqldrv
181 beairdragd aqvvtgylp lvsagdcpe1 gdvseadrw aveltgqine tvreaaerhd
241 alfvlpddad ehtscappqg zwadiagggt dayplhptsa gheamaaavr dalglepvqp

```

FIGURE 23 (SEQ ID No. 30)

ZP 00094165

```

1 mgqvklfarr capvllalag lapastvare aplaegaryv algssfaagp gvgnpaggsp
61 ercgrgtlvy phllaealkl dlvdatscga tthhvlgpwn evppqidsvn gdtlrvltti
121 ggndvsvfvg ifaaacekma spdprcgkwr eiteeewgad eermrsivrg iharaplrv
181 vvdvitytvp psqtcaamai spdrlaqsrs aakrlarita rvareegasl lkfshisrrh
241 hpcsakpwn qlsapaddgi pvhpnrlgha eaaaalvklv klmk //

```

FIGURE 24 (SEQ ID No. 31)

NP_625998.

```

1 mrrfrlvglf sslvlaagaa ltgaataqaa qpaaadgyva lgdsyssgvg agsyissgd
61 ckrstkahpy lwaaahspst fdrtacsgar tgdvlgqqlg plssgtglvs isiggndagf
121 adtmittcvlg sessclsrta taasyvdstl pgkldgvysa isdkapnahv wvigyprfyk
181 lgttciglse tkrtainkas dhlntvlaqr aaahgftfgd vrttftghel cagspwlhsv
241 nwlmgiesyh ptaagggggy lplvngaa
//

```

FIGURE 25 (SEQ ID No. 32)

NP_827753.

```

1 mrrsritayv tslllavgca ltgaataqas paaaatgyva lgdsyssgvg agsyissgd
61 ckrsskaypy lwqaahspss fsfnacsgar tgdvlanglg tlnestglvs ltiggndagf
121 sdvmttcvlg sdsaclsrin takayvdstl pgqldsvyta istkapsahv avlgyprfyk
181 lggsclagls etkrsainda adylnsaiak readhgftfg dvkstftghe icssstwlhs
241 ldllnigqsy hptaagqsgg ylpvmnsva
//

```

FIGURE 26 (SEQ ID No. 33)

```

MRLTRSLSAASVIVFALLLALLGISPAQAAGPAYVALGDSYSSGNGAGSYIDSSGDCHRSN
NAYPARWAAANAPSSFTFAACSGAVTTDVINNQLGALNASTGLVSTITIGNDAGFADAMTT
CVTSSDSTCLNRLATATWYINTTLLARLDVYSQIKARAFNARVVVLGYPRMYLASNEWYC
LGLSNTKRAAINTTACTLNSVSSRATAGFRFGDVRPTNNHELFPENDWLHSLTLPVWE
SYHPTSTGHQSGYLPVLNANSST

```

FIGURE 27 (SEQ ID No. 34)

```

ADSRPAFSRI VMFGDSLSDT GKMYSKMRGY LPSSPPYYEG RFSNGPVWLE QLTKEPGLT
AVAYNKISWNPKYQVINNLDYEVTFQLQKDSFKPDDLVLWVGANDYLAYGWNTQDAKVRDAISDAAN
RMVLNGAKEILLFNLPDLGQNPSARSQKVVEAASHVSAYHNQLLLNLARQLAPTGMVKLFEIDKQFAEML
RDPQNFGLSDQRNACYGGSYVWKPFAASRSASTDSQLSAFNPQERLAIAGNPLLAQAVASPMARSASTLN
CE
GKMFWDQVHPTTVVHAALSEPAATFIESQYEFIAH

```

FIGURE 28 (SEQ ID No. 35)

```

1 ADTRPAFSRI VMFGDSLSDT GKMYSKMRGY LPSSPPYYEG RFSNGPVWLE QLTKEPGLT
61 IANEAEAGGAT AVAYNKISWN PKYQVINNLD YEVTFQLQKD SFKPDDLVLW VGANDYLAY
121 GWNTQDAKR VRDAISDAAN RMVLNGAKQI LLFNLPDLGQ NPSARSQKVV EAVSHVSAYH
181 NKLLNLARQ LAPTEMVKLF EIDKQFAEML RDPQNFGLSD VENPCYDGGY VWKPFATRSV
241 STDRQLSAFS PQERLAIAGN PLLAQAVASP MARRSASPLN CEKMFWDQV HPTTVVHAAL
301 SERAATFIET QYEFIAHG

```

FIGURE 29 (SEQ ID No. 36)

ACAGGCCGATGCACGGPACCGTACCTTTCCGCAGTGAAGCGCTCTCCCCCATCGTTCCG
 CCGGACTTCATCCGCGATTTTGGCATGAACACTTCCTTCAACGCGGTAGCTTGCTACAA
 GTGCGGCAGCAGACCGCTCGTTGGAGGCTCAGTGAGATTCAACCGATCCCTGTGGGCCG
 CATCGCTCATCGTCTTCGCGCTGCTGCTCGCGCTGCTGGGCATCAGCCCGGCCAGGCAG
 CCGGCCCGGCCTAIGTGGCCCTGGGGGATTCCCTATTCTCGGGCAACGGCGCCGGAAGTT
 ACATCGATTTCGAGCGGTGACTGTACCGCAGCAACAACGCTACCCCGCCCGCTGGGCGG
 CGGCCAACGCACCGTCCCTTCACCTTCGCGGCTCGCTCGGGAGCGGTGACCACGGATG
 TGATCAACAATCAGCTGGGCGCCCTCAACGCGTCCACCGGCTGCTGAGCATCACCATCG
 GCGGCAATGACGCGGCTTCGCGGACGGGATGACCGCTGCGCTCACGAGCTCGGACAGCA
 CCTGCTCAACCGGCTGSCACCGCCACCACTACATCAACACCGCTGCTCGCCCGGC
 TCGACGCGGTCTACAGCCAGATCAAGGCCGTGCCCCCAACGCCCGCTGCTGCTCCTCG
 GCTACCCGCGCATGTACCTGSCCTCGAACCCCTGGTACTGCTGGGCTGAGCAACACCA
 AGCGCCGCGCCATCAACACACCGCCGACACCTCAACTCGGTGATCTCCTCCCGGCCA
 CGGCCACCGATTCCGATTTCGGCGATGTCCGCCCGACCTTCAACAACACGAACTGTTCCT
 TCGGCAACGACGCGTGCCTCACTCACCCTGCGCGTGTGGGAGTCTACACCCACCA
 GCACGGGCCATCAGAGCGGCTATCTGCGGTCCTCAACGCCACAGCTCGACCTGATCAA
 CGCACGGCGGTGCCCCCGCGCGCTCAGGCTCGGCGCGGGCGCGCAGCGCGTTGATCA
 GCCCACAGTGCCTGTGACGGTCCCACCTCAAGGTTCGAGGTTGTACGTACGCTGCGGCC
 GCTCCAGAAGTGGAACTCAGCAGGACCGTGGAGCCGTCCCTGACCTCGTCAAGAACTC
 CGGGGTCAAGCTGATCACCCCTCCCCGTAGCCGGGGGGAAGGCGGGGCCCAACTCCTT
 GTAGGACCTCCAGTCTGCGGCCCGGCTTGCCACCGTCCGCGTAGACCGCTTCCATGCT
 CGCCAGCCGGTCCCCGCGGAACCTCGGTGGGATGTCCGTGCCCAAGGTGGTCCCGGTGT
 GTCCGAGAGCACCGGGGGCTCGTACCGGATGATGTGCAGATCCAAAGAAAT

FIGURE 30 (SEQ ID NO. 37):

MRLTRLSAASVIVFALLALLGISPAQAAGPAYVALGDSYSSGNGAGSSYIDSSGDCHRSN
 NAYPARWAANAAPSSFTFACSGAVTTDVINNQLGALNASTGLVSTITGGNDAGFADAMTT
 CVTSSDSTCLNRLATATNYINTILLARLDAVYSQIKARAPNARVVVLGYPRMYLASNPWYC
 LGLSNTKRAAINTTADTLNSVISRATAHGFREFDVRPTFNNHELFFGNDWLHSLTLPVWE
 SYEPTSTGHQSGYLFVLNANSST

FIGURE 31 (SEQ ID No. 38)

```

1 mlphpagerg avgaiffallv gtpgdrrrlrl echetrlplrg rcgcgerrvp pltlpgdgvl
61 cttsstrndae tvwrkhlqpr pdggfrphlg vgciliagqgs pgvlwogreg crfevcrddt
121 pglstrtrngd ssppfragws lppkcgelsq sarktpavpr yslrltdrpd gprgrfvsgs
181 praatrrrlf lgipaivlvt altlvlavpt gretlwmwc eatqdwclgv pvdrrggpae
241 dgefilllspv gaatwgnyya lgdsyssgdg ardyypgtav kggcwrsana ypelvaeayd
301 faghlsflac sgqrgyamld aidevgsqld wnsphslvt igiggnldgf stvlktcmvr
361 vpildskact dgedairkrm akfettfeel isevrtrapd azilvvgypr ifpeeptgay
421 ytitasnqrw lnetiqefnq qlaaavavhd eeiaasgcvg svefvdvyha ldgheigsde
481 pwnvgvqlrd latgvtvdrs tfhpneaghr avgervieqi etgpggplya tfavvagatv
541 dtlagevg

```

FIGURE 32 (SEQ ID No. 39)

```

1 ggtggtgaac cagaacaccc ggtcgtcgcc gtgggctcc aggtgcaggt gcaggttctt
61 caactgctcc agcaggatgc cgcggtggcc gtgcacgatg gccttgggca ggctgtggt
121 ccccgacgag tacagcaccc atagcggatg gtgcacacgc agcggggtga actccagttc
181 cgcgccttcg cccgcggctt cgaactccgc ccaggacagg gtgtcggcga caggcccgca
241 gccaggtac gccaggacga cgggtgtgtg caggctgggc atgcgctgc gcagggcttt
301 gagcacgtca cggcggtcga agtctctacc gccgtagcgg tagccgtcca cggccagcag
361 cactttcgtt tcgactctcg cgaaccggtc gaggacgctg cgaaccctga agtcggggga
421 acaggacgac caggctcgac cgatcgcgcc gcaggcgagg aatgcggccg tcgctcggc
481 gatgttcggc aggtagggca cgaaccggtc gccggggccc accccgaggc tgcggagggc
541 cgcagcgatc cggcggtgc gggctccgag ttctccccag gtccactcgg tcaacggccg
601 gaggctcgac gcgtgcccga tcccacggc tgatgggtca cggctcgcca agatgtgtc
661 gccgtagtgt aggggtggcc cggggaaaca gacggcgccg gccatggcgt cggaggcgag
721 tcctgtggtg taagggtgga cggcgcgcac ccggtagtac tccagatcg cggaccagaa
781 tccttcgagg tcggttaccc accagcgcca cagtgcctcg tagtcgggtg cgtccacacc
841 gcggtgctcc cgcaccacgc ggggtgaacg ggtgaggttg gccgttctt tgcgtcctc
901 gtccggactc cacaggatgc gccgctgcgg cttagtgtc atgaaacgcg accctgtcgt
961 ggacggtgag gatgcgggtg gcgtcgggtg cctccctaa cgtcccccgg tgcaggagt
1021 ttgtgcacca catctagcac gccggacgcg gaaaccgtat ggagaaaaa cctacaaccc
1081 cggcgagacg gtgggttttc gccacactta ggggtcgggt gccctgcttg cctcagggga
1141 agtcccgggg tgctgtggtg cggcggggag ggctgtcgtc tcgaggtgtg cggcggggac
1201 actccgggac tcagccgtac ccgcaacggg gacagtcttc ctcccttcgg ggtggatgg
1261 tccttcctcc cgaactgcgg cagatcttcc cagtacgccc ggaaaaaacc cgtgtgcgcc
1321 aggtactctt tcttccgaac agacaggccg gacggtccac gggggagggt tgtgggcagc
1381 ggaccacggt cggcgaccag acgacggttg ttctcgggta tcccgcgtct tgtacttgt
1441 acagcgctca cgtgtgtctt ggtgtcccg acggggcgcg agacgctgtg gccatgttg
1501 tgtgaggcca cccaggactg gtgcctgggg gtgcgggtcg actcccgcg acagcctgcg
1561 gaggacggcg agttctgtct gttctctccg gtccagggag cgaactgggg gaactattac
1621 gcgctcgggg attcgtactc ttccggggac gggggcccgc actactatcc gccacccgcg
1681 gtgaaggggc gttgctgycg gtccgctaac gcctatccgg agctggtcgc cgaagcctac
1741 gacttcgccc gacacttgtc gttctcggcc tgcagcgccc agcgcggtca cgcctgtct
1801 gacgctatcg acgaggtcgg ctccgagctg gactggaaat cccctccac gtctgtgtg
1861 acgatcggga tcggcgccaa cgatctgggg ttctccacgg ttttgaagac ctgcatgtg
1921 cgggtgcggc tgcctggacg caaggcgtgc acggaccagg aggaagctat ccgcaagcgg
1981 atggcgaat tcgagacgac gtttgaagag ctcatcagcg aactgcgcac ccgcgcgcg
2041 gacgcccga tcttctgtgt gggctacccc cggatcttcc cggaggaaac gaccggcgcc
2101 tactacacgc tgaccgcgag caaccagcgg tggctcaacg aaaccattca ggagtccaac
2161 cagcagctcg ccgaggtcgt cgcggtccac gacgaggaga ttgcgcgctc gggcggtgtg
2221 ggcagcgtgg agttcgtgga cgtctaccac gcgttggaag gccacgggat cggctcggac
2281 gagcgtggg tgaacggggt gcagttcggg gacctcgcca cgggggtgac tgtggaacgc
2341 agtaccttcc accccaacgc cgtcgggcac cggcggtcgt gtagcggtt catcgagcag
2401 atcgaacccg gcccgggcgg tccgctctat gccactttcg cgggtgtggc gggggcgacc
2461 gtggacactc tcgcgggcga ggtgggggta cccggtctac cgtccggccc gcaggtctgc
2521 gagcactgcg gcgactgtgt ccaactgccca gtgcagttcg tcttcgggtg tcaccaggg
2581 cggggagagc cggatcgttg agcgtgctgt gtctttgacg agcacacccc gctgcaggag
2641 ccgttcgcac agttctcttc cgttgccacg agtcgggtcg acgtogatcc cagcccaacg
2701 gccgatgctg cggggcgcca ccacgcccgt gccgaccagt tggtcgaggg gggcgcgag
2761 caccggggcg agggcgcgga catggtccag gtaaggggcg tcggcgagga ggctcaccac
2821 ggcagtcggc accgcgcagg cagggcgctt gccgcgaag gtgctgacct gctggcggg
2881 gcggtcacg tcgaagactt ccgcgtcgcc taaccggccc gccacgggca ggtgcgcgc
2941 gccagcgct ttgcgaaca ggtagatata ggcgtcgact ccgctgiggt cgcaggcccg

```

FIGURE 33 (SEQ ID No. 40)

```

1  vsggpraatr  rrlflgipal  vlvttaltlwl  avptgretlw  rrwceatqdw  clgvpvdsrg
61  qpaedgefil  lspvqaatwg  nyyalgdsys  sgdgardyyp  gtavkggcwr  sanaypelva
121 eaydfaghl  flacsggrgy  amldaiderg  sqldvmspht  slvtigiggn  dlglstvlkt
181 cmvrvplld  kactdgedai  rkrmakfett  feelisevrt  rapdariliv  gyprifpeep
241 tgayytilas  ngrwlnetiq  efnqqlaeav  avhdeeiias  ggvgsvafvd  vyhaloghei
301 gsdepwvngv  qlrdlatgvt  vdrstfhpna  aghravgerv  ieqietgpgp  plyatfavva
361 gatvdtlage  vg

```

FIGURE 34 (SEQ ID No. 41)

```

1  mrttviaasa  lllagcadg  areetagapp  gesseggiree  gaeastsitd  vyialgdsya
61  anggrdqplr  gepfclrssg  nypellhaev  tdtccgavt  gdllprtlg  ertlpaqvda
121 ltedttlvtl  siggndlgfg  evagcirari  agenaddcvd  llgetigqql  dqlppqldrv
181 heairdragd  eqvvvtgylp  lvsagdcpe  gdvseadrw  aveltgqine  tvreaaerhd
241 alfvlpddad  ehtscappqq  zwadiaggqt  daypihptsa  gheamaasvr  dalglepvqp

```

FIGURE 35 (SEQ ID No. 42)

```

1  ttctgggggtg ttatgggggtt gttatcggct cgtcctgggt ggatcccgcc aggtggggta
61  ttacacggggg actttttgtgt ocaacagccg agaatgagtg ccctgagcgg tgggaatgag
121  gtggggggggg ctgtgtcgcc atgagggggc ggcgggctct gtggtgcccc gcgacccccg
181  gccccgggtga gcggtgaatg aaatccggct gtaatcagca tcccgtgccc accccgtcgg
241  ggaggtcagc gcccgagtg tctacgcagt cggatcctct cggactcggc catgctgtcg
301  gcagcatcgc gctccgggt ctggcgctcc ctggctgtt ctgctgtctg tccctggag
361  gcgaatgat caccgggag tgatcacccg gtggtctcat cccggatgcc cacttcggcg
421  ccacccggca attcggggcag ctccgggttg aagtaggtgg catccgatgc gtccgtgacg
481  ccatagtggg cgaagatctc atcctgctcg aggtgtctca ggccactctc cggatcgata
541  tcgggggct ccttgatggc gtcctgtctg aaacggaggt gcagcttggt ggctaccaat
601  ttgcacccac ggagcgggac gaggctggaa tgacggccga agagcccggt gtggacctca
661  acgaaggtgg gtatcccggt gtcactattg aggaacacgc cctccacccg acccagcttg
721  tggccggagt tgcctaggc gctggcatcc agaaggga aaatctcata tttgtcggtg
781  tgctcagaca tgatcttct tgcctgtcgg tgctcggtag taaccaggta ggggtgaatg
841  caactgttat tttctgtta ttttaggaat tgggtccatc cccacaggct ggctgtggtc
901  aaatcgtcat caagtaatcc ctgtcacaca aaatgggtgg tgggaagccct ggtcgcggtt
961  ccgtgggggg cgcctgtccc cgcaggatcg tcggcatcgg cggatctggc cggtaecccg
1021  cggtagaata aatcattctg taaccttcat caogtttggt tttaggtatc cgcctcttc
1081  gtcctgaccc cgtcccccgc gcgcgggagc ccgcgggttg cggtagacag gggagacgtg
1141  gacacccatga ggacaacggt catcgacgca agcgcattac tcttctctgc cggatgcgcg
1201  gatggggggc gggaggagac cgcgggtgca ccgcgggtg agtccctccg gggcatccgg
1261  gaggaggggg cggaggcgct gacaagcatc accgacgtct acatccgccc cggggattcc
1321  tatgcgggga tggggggggc ggatcagccg ttacgggggt agccgttctg cctgcgctcg
1381  tccggttaatt acccggaact cctccacgca gaggtcaccc atctcacctg ccagggggcg
1441  gtgacccggg atctgctcga acccaggagc ctgggggagc gcaagctgac ggcgcaggtg
1501  gatgcgctga cggaggacac caccctgttc accctctcca tcgggggcaa tgacctcgga
1561  ttccggggagg tggcgggatg catccgggaa cggatcgccg gggagaaogc tgatgattgc
1621  gtgacccctg tgggggaaac catcggggag cagctcgatc agcttcccc gcagctggac
1681  cgcgtgcacg aggcctatcc ggaacgcgcc ggggacgcgc aggttggtgt caccggttac
1741  ctgcgcctcg tgtctgcgg ggaactgccc gaactggggg atgtctccga ggcggatcgt
1801  cgttggggcg ttgagctgac cgggcagatc aacgagaccg tcgcgcaggc ggcggaacga
1861  caogatgccc totttgtctt gcccgacgat gccgatgagc acaccagttg tgcaaccnca
1921  cagcagcgct gggcgatat ccagggccaa cagaccgatg cctatcugct gcacccgacc
1981  tccgcgggcc atgaggcgat ggccgcccgc gtcgggagc cgttgggccc ggaacgggtc
2041  cagccgtagc gccggggcg cgttctgtga cgaaccaacc atgcccaggt gcagtcacat
2101  ccgcacatag cgcgcgcggg cgaaggagta cgcacccatg aggatgagcc cgaatgcgac
2161  gatgatgagc agcacactgc cgaagggttg tccccgagg gtgcgcagag ccgagtcacg
2221  acctgcggcc tgcctccgat catgggccc aocggcgatg acgatcaaca ccccaggat
2281  ccggaaggcg ataccacggg cgaataaacc ggcgtgtccg gtgatgatga tccgggtccc
2341  gacctgccct gaccccgac ccgcctccag atcctcccg aaatcccggt tggcccccctt
2401  ccagaggttg tagacacccg cccccagtag caccagccc gcgaccacaa ccagacacc
2461  accccagggt tgggatagga cgggtggcgt gacatcggtg gcggtctccc catcgagggt
2521  gctgccgccc cgggcgaagg tggaggtggt caocccagg gagaagtaga ccattggccat
2581  gaccgcccc ttggcccttt ccttgaggtc ctgcgcgcgc agcagctggt tcaattgcca
2641  ggtgccagg gccgcaggg ccatgacggc aacccacagg aggaactgcc caccggagc
2701  ctccgcatg gtggccagg cactgaatt cgaagcctca tcaaccgaac cgcggatcc
2761  agtggcgatg cgaacggcga tcaacccgat gaggatgtgc agtatgccca ggacaatgaa
2821  accacctctg gccagggtgg tcagcgcggt ggggtcctcg gcctggtcgg cagcccggtc
2881  gatcgtccgt ttccgggato tgggtgtcgc cttatccata gctccattg aaccgccttg
2941  aggggtgggc ggcacatgta agggcggtt gtgatctgaa ctgtgatgtt ccatacaacc

```

FIGURE 36 (SEQ ID No. 43)

```

1 mrrfrlvglf sslvlaagaa ltgaataqaa qpaaadgyva lgdsysagvg agcyisssgd
61 ckrstkahpy lwaahspst fdftacsgar tgdvlsqqlg plssgtglvs isiggnadagf
121 admtttcvlq sssslsria taesyvdstl pgklidgyva isdkapnahv vvigypxfyk
181 lgttciglse tkrtainkas dhlnvlagr aaahgftfgd vrttftghel csqspwlhsv
241 nwlmgesyh ptaagqsggy lpvlngea

```

FIGURE 37 (SEQ ID No. 44)

```

1 ccggcgggcc cgtgcaggag cagcagccgg ccggcgatgt cctcgggcgt cgtcttcate
61 aggcgcgcca tctgctcgcc gaccggcgcc gtgtagttag ccgggaccto gtcccaggtg
121 ccggcgggca tctggcggtt ggtgcggtgc gggcgcgccc gaggggagac gtaaccagaag
181 cccatcgtea cgttctccgg ctgctggttc ggctcgtccg ccgtccgtc cgtcgcctcg
241 ccgagcaact tctcgcgag gtcggcgctg gtcccgctca ccgtgacgtc ggcccccgg
301 ctccagcgcg agatcagcag cgtccagcgg tccgactccg ccagcgtcgc gctgcggtcg
361 tctgctcgcc cgatccgcag cccggcgccg ccggcgcgca gcagcgtggc gccggaccgt
421 acgcggtcga tgttcgcccc gtgcgagtag ggctgctcac ccgtggcgaa accggcgagg
481 aacagcgcggt cgcgcagtc ggacggggag tccgtgtcgt ccaagttag ccggatcgcc
541 agggcttctg cggggttcac ggacatgtcg ccatgatcgg gcacccggcc gcgcgctgca
601 ccgcttctcc cgggcacgca cgaacggggc ttctcgccg tcttcgtcc gaacttgaac
661 gagtgtcagc ctttcttgg catggacact tccagtcaac gcgcgtagct gctaccacgg
721 ttgtggcagc aatctcgcta agggaggttc catgagacgt ttccgaattg tccgttctct
781 gagtctgctc gtcttcgccc ccggcgccgc cctcaccggg gcagcgaccg cccagcgggc
841 ccaaccggcc gccgcgagc gctatgtggc cctcgggcag tctactact cccgggtcgg
901 agcgggcagc tacatcagct cgaagcgcca ctgcagagcc agcaggaagg cccatcccta
961 cctgtggggc gccgccact cgcctccac gttcgacttc accgcctgtt ccggcgcccg
1021 tccgggtgat gttctctccg gacagctcgg ccggctcagc tccggcaccg gctcgtctc
1081 gatcagcacc ggccgcaacg acccggtttt cgcgacacc atgacgacct gtgtgtcca
1141 gtccgagagc tctgtcctgt ccgggatcgc caccgcccag gcgtacgtcg actcgacgt
1201 gcccgccaag ctgcagcgcc tctactcggc aatcagcgac aaggcgccga accgccacgt
1261 cgtcgtcacc ggctaccgcc gttctacaa gctcggcacc acctgcctcg gctgtccga
1321 gaccagcgcg accggcgatca acaaggcctc cgaaccctc aacacgtcc tcccccagg
1381 cgcggccgcc cagcgcttca ccttcggcga cgtacgcacc accttcaccg gccacgagct
1441 gtgctccggc agcccttggc tgcacagcgt caactggctg aacatcgccg agtcgtacca
1501 cccaccggcg gccggccagt ccggtggcta cctgcgggtc ctcaacggcg ccgcttgacc
1561 tcaggcgaaa ggagcagaa gggagcgga gggagacgag gagtgggagg ccccgccga
1621 cggggtcccc gtcccgctct cgtctcgt ccggtcccg caagtccag agaacggcac
1681 cgcgtcggac gtggcccgca ccggactccg cactccacg cgcacggcac tctcgaacgc
1741 gccggtgtcg tctgtcgtcg tcaccaccac gccgtcctgg cgcgagcgt ccgcccga
1801 cgggaaggac agcgtccgcc accccggatc ggagaccgac ccgtcccgcg tcaaccaccg
1861 gtacccgacc tccgcccga gccgccgac cgtgaacgtc gccgtgaacg cgggtgcccg
1921 gtcgtcggc ggcgacagg ccccagata gtgggtgcgc _gagctcacca cggtaacctc
1981 caccgactgc gctcggggc

```

FIGURE 38 (SEQ ID No. 45)

```

1 mrrsrityyv tslllavgca ltgaataqas paaaatgyva lgdsysagvg agsylessgd
61 ckrsskaypy lwqaahspss fsfmacsgar tgdvlanglg tlnsstglvs ltiggnndagf
121 sdvmttcvlg sdsaclsrin takayvdstl pggldsvyta istkapsahv avlgyprfyk
181 lggsclagls etkrseinda adylnsaiak raadhgftfg dvkstftghe icsstwlhs
241 ldlilnigqsy hptaaggsgg ylpvmnsva

```

FIGURE 39 (SEQ ID No. 46)

```

1 ccaccgcccgt gtcggcgccgt agtctcctgg cctcggtcgc ggagaggttg gcogtqtage
61 cgttcagcgc ggccgcgaac gtcttcttca ccgtgccgcgc gtactcgttg atcaggccct
121 tgccttgct cgacgcggcc ttgaagccgg tgccttctt gagcgtgacg atgtagctgc
181 ccttgatcgc ggtgggggag ccggcgccga gcaccgtgcc ctccggccgg gtggcctggg
241 cgggcagtcg ggtgaatccg cccacgaggg ccgcggtcgc caaggcggtt atcgcgcca
301 tccggatctt ctgtctacgc agctgtgcca tacgagggag tctcctctg ggacggcgcc
361 cgcctgggtg gggcgccagg ctgtgggggg tgcgcgcgc atcacgcaca cggccctgga
421 gcgtcgtgtt ccgccctggg ttgagtaaag cctcgcccat ctacgggggt ggctcaagg
481 agttgagacc ctgtcatgag tctgacatga gcacgcacac aacggggccg tgagcaccoc
541 ggggcgaccc cggaaagtgc cgagaagtct tggcatggac acttcctgtc aacacgcgt
601 gctgttacga cggttacggc agagatcctg ctaaaggag gttccatgag acgttccga
661 attacggcat acgtgacctc actcctctc gccgtcggct gcgccctac cggggcagcg
721 acggcgccag cgtcccccgc cgcgcggcc acgggctatg tggccctgg cgactcgtac
781 tcgtccygtg tcggcgccgg cagctacctc agctccagcg gcgactgcaa ggcaggttcg
841 aaggcctatc cgtacctctg gcaggcccgcc cattcacctc cgtcgttcag ttccatggct
901 tgctcggggc ctcgtaaggg tgatgtcctg gccaatcagc tcggcaccct gaactcgtcc
961 accggccctg tctccctcac catcggaggg aacgacgcgg gcttctccga cgtcatgacg
1021 acctgtgtgc tccagtcoga cagcgccctg ctctcccgca tcaacacggc gaagcgctac
1081 gtcgactcca cctgcccgg ccaactcgac agcgtgtaca cggcgatcag cccgaaggcc
1141 ccgtcggccc atgtggccgt gctgggctac ccccgcttct acaactggg cggctcctgc
1201 ctgcggggcc tctcggagac caagcggctc gccatcaacg acgcgccga ctatctgaac
1261 agcgccatcg ccaagcgccg cgcgcaccac ggttcacctc tcggcgacgt caagagcacc
1321 ttcacggcc atgagatctg ctccagcagc acctggctgc acagtctga cctgtgaac
1381 atcgccaggt ctaccacccc gaccggggcc ggccagtcgc gcggtatct gccggtcatg
1441 aacagcgtgg cctgagctcc pacggccctga atttttaagg cctgaatttt taaggcgagg
1501 gtgaecogga agcggaggcc ccgtccgtcg gggctctcgt cgcacaggtc aocgagacg
1561 gcacggagtt ggaagtcgtg cgcacccggg cgcgcacctc gacggcgatc tcgttcgaga
1621 tcgttcggct cgtgtcgtac gtggtgacga acacctgctt ctgctgggtc tttccggccg
1681 tcgcccggaa ggacagctgc ttccagcccg gatccgggac ctgcgccctt ttggtacccc
1741 agcggtaact cacctcgacc ggcacccggc ccaccgtaga ggtcgcggtg aacgtggggc
1801 cctggggcgt gggggggggg caggcaccgg agtagtcggt gtgcacggcg gtgaccgtca
1861 ccttcacgga ctgggcgggc ggggtcgtcg tacgcgcgcc gccacccggc cctcccgag
1921 tggagcccca gctgtggtcg ccccgcgct cggcggtgtc gtccctgggg gttttcgaa

```

FIGURE 40 (SEQ ID No. 47)

```

1 mgsqpraatz rrlflgipal vltaltlvtl avptgretlw mmwceatqdw clgvpvdsrg
61 qpaedgefll lspvqaatwg nyvalgdsys sqdgardyyp gtavkkgcwr sanaypelva
121 eaydfaghls flacsgqrgy amldaidevg sqldwnspht slvtigiggn dlqfstvlkt
181 cmvrvpllds kactdgadai rkrmakfett feelisevrt rapdarilvv gyprifpcep
241 tgayytltas nqrwlnetiq efnqqlaeav avhdeeiias ggvgsvefvd vyhaldghei
301 gsdepwngv qlrdlatgvt vdrstfhpna aghravgerv ieqiatgpgg plyatfavva
361 gatvdtlage vg

```

FIGURE 41 (SEQ ID No. 48)

```

1      ctgcagacac ccgccccgcc ttctcccgga tctcatgtt cggcgactoc ctcagcgaca
61      cgggcaagat gtactccaag atgcgcggct acctgcccgc ctccccgcgc tactacgagg
121     gccgcttctc gaacggcccg gtctggctgg agcagctgac gaagcagttc cccggcctga
181     cgatcgccaa cgaggccgag gggggcgcgga ccgcagtcgc ctacaacaag atctcctgga
241     acccgaagta ccaggtcatt aacaacctcg actccgaggt caccagttc ttgcagaagg
301     actcggtcaa gcccgacgac ctgggtcatcc tgtgggtggg cgcacaacgac tacctggcct
361     acggttgaa caccggagcag gacgccaagc ggtgcgcgga cgcctctcg gacgcggcaa
421     accgcatggt cctgaacggc gcgaagcaga tcttgcgtgt caacctgccc gacctgggcc
481     agaaccgctc cgcctcgctc cagaaggctc tgcaggccgt ctgcgacgtg tccgcctacc
541     acaacaagct gctcctcaac ctgcgccggc agctcgcccc gacgggcatg gtcaagctgt
601     tcgagatoga caagcagttc gcggagatgc tgcgcgaccc ccagaacttc ggcttgagcg
661     aegtggagaa cccgtgctac gacggcggtc acgtgtgaa gccgttcgcc acccggtccg
721     tctgaccga ccggcagctg tggccttct cgcgccaggga gcgctggcg atcgctggca
781     accgctcct gccacaggcg gtacgttcgc cgtatggcccg ccgctcgcc tcgcccctca
841     actgcgaggg caagatgttc tgggaccagg tccacccac caccgtggtc caccgcgcc
901     tctcgagcg cgcgccacc ttcatcgaga cccagtaga gtctctcgcc cactagtcta
961     gaggatcc

```

FIGURE 42

1. L131
2. *S. avermitilis*
3. *T. fusca*
4. Consensus

```

      1                                     50
1 (1) -----MRLTRSLAASVIVFALLLALLGISPAQAAG-----
2 (1) -----MRRSRITAYVTSLLLAVGCALTGAATAQASPA-----
3 (1) VGSGPRAATRRRLFLGIPALVLVTALTVLAVPTGRETLWRMWCEATQDW
4 (1)      MRRSRFLA ALILLTLA AL GAA ARAAP

      51                                     100
1 (32) -----P-AYVALGDSYSSGNGAGSYID
2 (33) -----AAATGYVALGDSYSSGVCAGSYLS
3 (51) CLGVFVDSRGQPAEDGEFLLLSPVQAATWGNYYALGDSYSSGDGARDYYP
4 (51)      A A YVALGDSYSSG GAGSY

      101                                    150
1 (53) SSGD---CHRSNNAYPARWAAANAP---SSFTFAACSGAVTTDVIN---
2 (57) SSGD---CKRSSKAYPYLWQAAHSP---SSFSEMACSGARTGDVLA---
3 (101) GTAVKGGCWRSANAYPELVAEAYDFA--GHLSFLACSGQRGYAMLDAIDE
4 (101) SSGD  C RSTKAYPALWAAAAH  SSFSF ACSGARTYDVLA

      151                                    200
1 (93) --NQLGALNAST--GLVSIITIGGNDAGTFADAMTTCVTS-----SDSTCL
2 (97) --NQLGTLNSST--GLVSLTIGGNDAGFSVDMTTCVLQ-----SDSACL
3 (149) VGSQLDWNSPRT--SLVTIGIGGNDLGFSTVLKTCMVR-----VPLLDS
4 (151)  QL LNS T  LVSITIGGNDAGFAD MTTCVL  SDSACL

      201                                    250
1 (133) NRLATATNYINTTLA-----RLDAVISQIKARAPNARVVVLGYPRMY
2 (137) SRINTAKAYVDSTLPG-----QLDSVYTAISTKAPSAHVAVLGYPRFY
3 (191) KACTDQEDAIRKMAKF---ETTFEELISEVRTRAPDARILVVGYPRIIF
4 (201) RIA AK YI  TLPA  RLDSVYSAI TRAP ARVVVLGYPRIY

      251                                    300
1 (176) LASNPWYCLGLSNTKRAAINTTADTLNSVISSRATAH-----GF
2 (180) KLGG-SCLAGLSETKRSAINDAADYLSAIAKRAADH-----GF
3 (237) PEEPTGAYYTLTASNQRWLNETIQEFNQQLAEAVAVHDEEIAASGGVGSV
4 (251)  SG  LGLS TKRAAINDAAD LNSVIKRAADH  GF

      301                                    350
1 (215) RFGDVRPTFNNHELFFGNDWLHSLTLP-----VWESYH
2 (218) TFGDVKSTFTGHEICSSSTWLHSLDLIN-----IGQSYH
3 (287) EFVDVYHALDGHEIGSDEPWVNGVQLRDLATC-----VTVDRSTFE
4 (301) TFGDV  TF GHELCSA PWLHSLTLP  V SYH

      351                                    395
1 (248) PTSTGHQSGYLPVLNANSST-----
2 (252) PTAAGQSGGYLPVMNSVA-----
3 (328) PNAAGHRAVGERVIEQIETGPGRPLYATFAVVAGATVDTIAGEVG
4 (351) PTA GHAAGYLEVLNSI T

```

FIGURE 43

SEQ ID No 17 which is the amino acid sequence of a lipid acyltransferase from *Candida parapsilosis*;

```

MRYFAIAFLI INTISAFVLA PKKPSQDDFY TPPOGYEAQ2 LGSILKTRNV PNPLTNVFTF VKVQNAWQLL
VRSEDTFGNP NAIVTIIQP FNAKKDKLVS YQTFEDSGKL DCAPSYAIQY GSDISTLTQ GEMYYISALL
DQGYVVVTPD YEGPKSTFTV GLQSGRATLN SLRATLKSGN LTGVSSDAET LLWGYSGGSL ASGWAAAIQK
EYAPELSKNL LGAALGGFVT NITATAEAVD SGPFAGIISN ALAGIGNEYF DFKNYLLKKV SPLLSITYRL
GNTHCLLDGG IAYFGKSFFS
RIIRYFPDGW DLVNQEPIKT ILQDNGLVYQ PKDLTPQIPL FIYHGTLDAL VPIVNSRKTFF QQWCDWGLKS
GEYNEDLTNG RITESIVGAP AALTWIINRF NGQPFVDGCQ HNVRASNLEY PGTPQSIKNY FEALHAILG
FDLGPDVKRD KVTLGGLLKL ERFAP

```

FIGURE 44

SEQ ID No 18 which is the amino acid sequence of a lipid acyltransferase from *Candida parapsilosis*;

```

MRYFAIAFLI INTISAFVLA PKKPSQDDFY TPPOGYEAQ2 LGSILKTRNV PNPLTNVFTF VKVQNAWQLL
VRSEDTFGNP NAIVTIIQP FNAKKDKLVS YQTFEDSGKL DCAPSYAIQY GSDISTLTQ GEMYYISALL
DQGYVVVTPD YEGPKSTFTV GLQSGRATLN SLRATLKSGN LTGVSSDAET LLWGYSGGSL ASGWAAAIQK
EYAPELSKNL LGAALGGFVT NITATAEAVD SGPFAGIISN ALAGIGNEYF DFKNYLLKKV SPLLSITYRL
GNTHCLLDGG IAYFGKSFFS RIIRYFPDGW DLVNQEPIKT ILQDNGLVYQ PKDLTPQIPL FIYHGTLDAL
VPIVNSRKTFF QQWCDWGLKS GEYNEDLTNG RITESIVGAP AALTWIINRF NGQPFVDGCQ HNVRASNLEY
PGTPQSIKNY FEALHAILG FDLGPDVKRD KVTLGGLLKL ERFAPHHHHE H

```

FIGURE 45

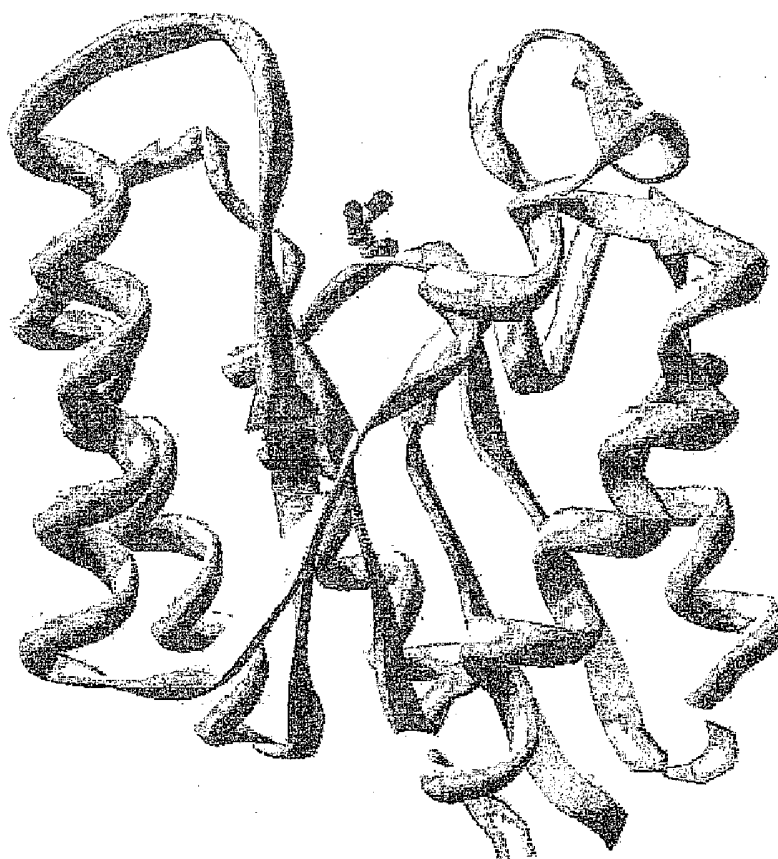


FIGURE 46

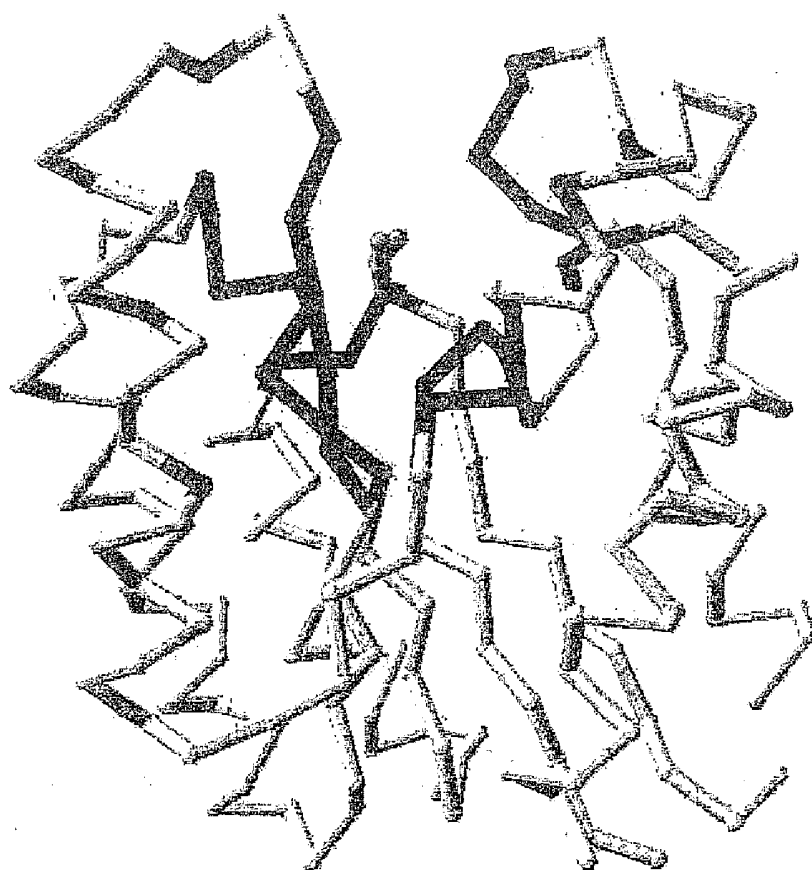


FIGURE 47

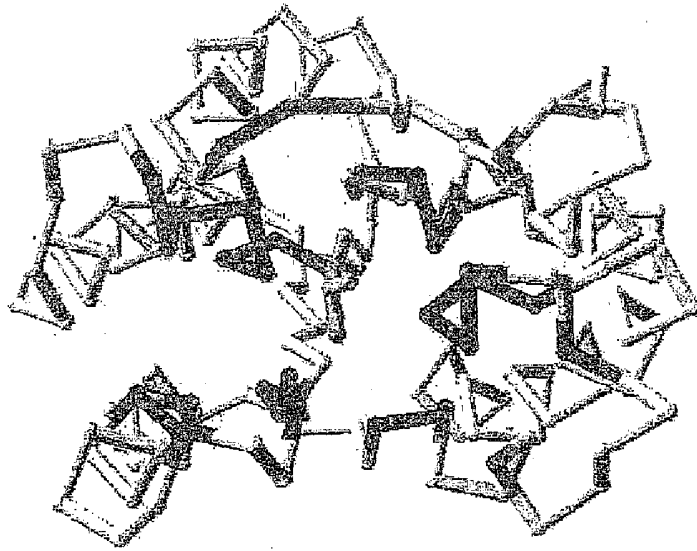


FIGURE 48

[illegible]

[illegible]

FIGURE 50

[illegible]

LD1DEOm	L T F P A	Y	Z E N A A K L F T A K	G A K V	I L S Q T P N N P	W E T G T F V N S P T N	
	h4h4h4h4	h4	h4h4h4h4h4h4h4h4h4	s5s5s5	s5 s5 s5		
LD1VNm	Q O T E Q	T	L R Q I L Q D V K a A	N A E P	I L m q i R L P A N	Y G R - - - - -	R Y N E A
	h4h4h4h4h4	h4	h4h4h4h4h4h4h4h4h4	s5s5s5	s5 s5 s5	s5s5s5s5s5s5	
LD1P1048Om	K R V R D	A	I S D A A N R R V L N	G A K E	I L L F N L P d l g	q n p S A R S Q K V V E A S H V S A	
LD1DEO	F V E Y A	E	L A A E V A - - - - -	- - - - -	- - - - -	G V E Y V D H W S Y V O S I Y E T	L G N A L V N
	h5h5h5h5h5	h5	h5h5h5h5h5h5h5	- - - - -	- - - - -	s6 s6 s6s6h6h6h6h6h6h6h6	h h h h h
LD1VNm	N E A F S	A	I Y P K L A k e - - - - -	- - - - -	- - - - -	F D V P L L P F M L E V Y L K P Q W - - - - -	
	h5h5h5h5h5	h5	h5h5h5h5h5h5h5h5	h5	h5	s6 s6 s6s6? h6h6h6h6h6	s
LD1P1048O	S A Y H N	Q	L L L N L A z q l a p	t g m v k l f e i	D K Q F A E N L R D P Q N F G L S D O R N a c o		
LD1DEOm	F V E Y A	E	L A A E V A - - - - -	- - - - -	- - - - -	G V E Y V D H W S Y V O S I Y E T	L G N A L V N
	h5h5h5h5h5	h5	h5h5h5h5h5h5h5h5	- - - - -	- - - - -	s6 s6 s6 s6h6h6 h6h6h6h6h6h6h6	h h h h h
LD1VNm	F S A I Y	P	K L A k e - - - - -	- - - - -	- - - - -	F D V P L L P F F M E E V Y	L K P Q W - - - - -
	h5h5h5h5h5	h5	h5h5h5h5h5h5h5h5	h5	h5	s6 s6 s6 s6? h6 h6 h6 h6h6	s
LD1P1048Om	Y H N Q L	L	N L N L A z q l a p	t g m v k l f e i	D K Q F A E N L R D P Q N F G L S D O R N a c y 99		
LD1DEO	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -
LD1VNm	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -
LD1P1048O	g s y v w	k p f a s i r s a s t d s q l s a f	n p q e x l a i a g n p l l a q a v a s p m a a r s a				
LD1DEOm	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -
LD1VNm	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -
LD1P1048Om	s y v w k p	f a s r s a s t d s q l s a f	n p q e x l a i a g n p l l a q a v a s p m a a r s a s t				
LD1DEO	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -
LD1VNm	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -
LD1P1048O	s t l n c e	g k M F W D Q V H P T V V V H A L S E P A A T F I E S O Y E F L A H -					
LD1DEOm	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -
LD1VNm	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -
LD1P1048Om	l n c e g	k M F W D Q V H P T V V V H A L S E P A A T F I E S O Y E F L A H -					

FIGURE 51

```

10      20      30      40      50
60
....*....|....*....|....*....|....*....|....*....|....*....|
1IVN_A      4 LLILGDSLSAG-----YRMSASAAWPALLNDKWqsk---
----- 34
P10480      28
IVMFGDSLSDTgkmyskmgylpssppyyeGRFENGFPVWLEQLTNEFPGLTianeeagpp 87

70      80      90      100     110
120
....*....|....*....|....*....|....*....|....*....|....*....|
1IVN_A      35 -tsvVNASISGDT-----
SQQCLARLPALLKQHQPWW 65
P10480      88 tavaYNKISWNPKyg-----
VINNLDYEVTQFLQKDSFKPDDL 125

130     140     150     160     170
180
....*....|....*....|....*....|....*....|....*....|....*....|
1IVN_A      66 VLVELGGNDG-----
LRGFQFPQQTET 87
P10480     126 VILWVGANDY-----LA--
YGNWTEQDAKRVDA 152

190     200     210     220     230
240
....*....|....*....|....*....|....*....|....*....|....*....|
1IVN_A      88 LRQILODVkaANAEPllmqiRLPANYGR-----
----- 115
P10480     153 ISDAANRMV-LNGAK-----EILLFNLPdlg-----
----qnp 180

250     260     270     280     290
300
....*....|....*....|....*....|....*....|....*....|....*....|
1IVN_A     116 -----RYNEAFSAIYPKLake-----
FDVPLLPFFME 142
P10480     181 SRSQKVVEAASHVSAVHNQLLLNLArqlaptg-----
mvklfeIDKQFAEMLRD 230

310     320     330     340     350
360
....*....|....*....|....*....|....*....|....*....|....*....|
1IVN_A     143 EVYLPQR-----
----- 150
P10480     231
PQNFGLSQQRNacyggsyvwkpfasrsastdsqIsafnpqerlaiagnpllaqavaspm 290

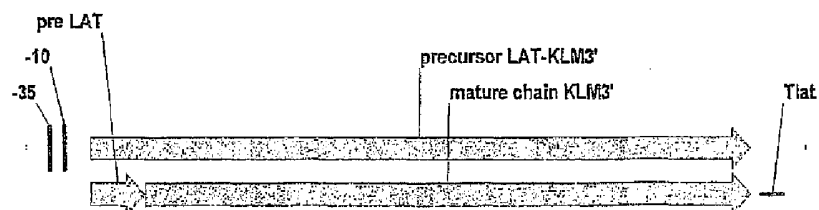
370     380     390     400
1IVN_A     151 -----MQDDGI-----HPNRDAQPFADWM 170
P10480     291 arsastlncegkMFWDQV-----HPTTVVHAALSEPA 322

```

FIGURE 52

		1	50
P10480	(1)	MKKWFVCLLGLVALTQQAADSRPAFSRIVMFGDLSLDTGKMYSKMRGYLP	
A. sal	(1)	-----ADTRPAFSRIVMFGDLSLDTGKMYSKMRGYLP	
A. hyd	(1)	-----ADSRPAFSRIVMFGDLSLDTGKMYSKMRGYLP	
Consensus	(1)	AD* RPAFSRIVMFGDLSLDTGKMYSKMRGYLP	
		51	100
P10480	(51)	SSPPYYEGRFSNGFPVWLEQLTNEFPGLTIANEAEGGPTAVAYNKISWNEK	
A. sal	(33)	SSPPYYEGRFSNGFPVWLEQLTKQFPGLTIANEAEGGATAVAYNKISWNEK	
A. hyd	(33)	SSPPYYEGRFSNGFPVWLEQLTKQFPGLTIANEAEGGATAVAYNKISWNEK	
Consensus	(51)	SSPPYYEGRFSNGFPVWLEQLT**FPGLTIANEAEGG* TAVAYNKISWNEK	
		101	150
P10480	(101)	YQVINNLDYEVTOFLQKDSFKPDDLVLWVGANDYLAYGWNTAQDAKVR	
A. sal	(83)	YQVINNLDYEVTOFLQKDSFKPDDLVLWVGANDYLAYGWNTAQDAKVR	
A. hyd	(83)	YQVINNLDYEVTOFLQKDSFKPDDLVLWVGANDYLAYGWNTAQDAKVR	
Consensus	(101)	YQVINNLDYEVTOFLQKDSFKPDDLVLWVGANDYLAYGWNTAQDAKVR	
		151	200
P10480	(151)	DAISDAANRMVLNGAKQILLFNLPDLGQNPSARSQKVVEAASHVSAYHNQ	
A. sal	(133)	DAISDAANRMVLNGAKQILLFNLPDLGQNPSARSQKVVEAVSEVSAHNK	
A. hyd	(133)	DAISDAANRMVLNGAKQILLFNLPDLGQNPSARSQKVVEAVSHVSAYHNQ	
Consensus	(151)	DAISDAANRMVLNGAK* ILLFNLPDLGQNPSARSQKVVEA* SHVSAYHN*	
		201	250
P10480	(201)	LLLNLARQLAPTMVKLFEIDKQFAEMLRDPQNFGLSDQRNACYGGSYVW	
A. sal	(183)	LLLNLARQLAPTMVKLFEIDKQFAEMLRDPQNFGLSDVENPCYDGGYVW	
A. hyd	(183)	LLLNLARQLAPTMVKLFEIDKQFAEMLRDPQNFGLSDVENPCYDGGYVW	
Consensus	(201)	LLLNLARQLAPTMVKLFEIDKQFAEMLRDPQNFGLSD**N* CY* G* YVW	
		251	300
P10480	(251)	KPFASRSASTDSQLSAFNPQERLAIAGNPLLAQAVASPMARSASTLNCE	
A. sal	(233)	KPFATRSVSTDRLSAFSPQERLAIAGNPLLAQAVASPMARRSASF LNCE	
A. hyd	(233)	KPFATRSVSTDRLSAFSPQERLAIAGNPLLAQAVASPMARRSASF LNCE	
Consensus	(251)	KPFA* RS* STD* QLSAF* PQERLAIAGNPLLAQAVASPM* RSAS* LNCE	
		301	336
P10480	(301)	GKMFWDOVHPPTTVVHAALSEPAATFTIESQYEF LAH-	
A. sal	(283)	GKMFWDOVHPPTTVVHAALSEAAATFTETQYEF LAHG	
A. hyd	(283)	GKMFWDOVHPPTTVVHAALSEAAATFTIANQYEF LAH-	
Consensus	(301)	GKMFWDOVHPPTTVVHAALSE* AATFTI** QYEF LAH*	

FIGURE 53



Gene construct for KLM3' expression

1225 bp

FIGURE 54

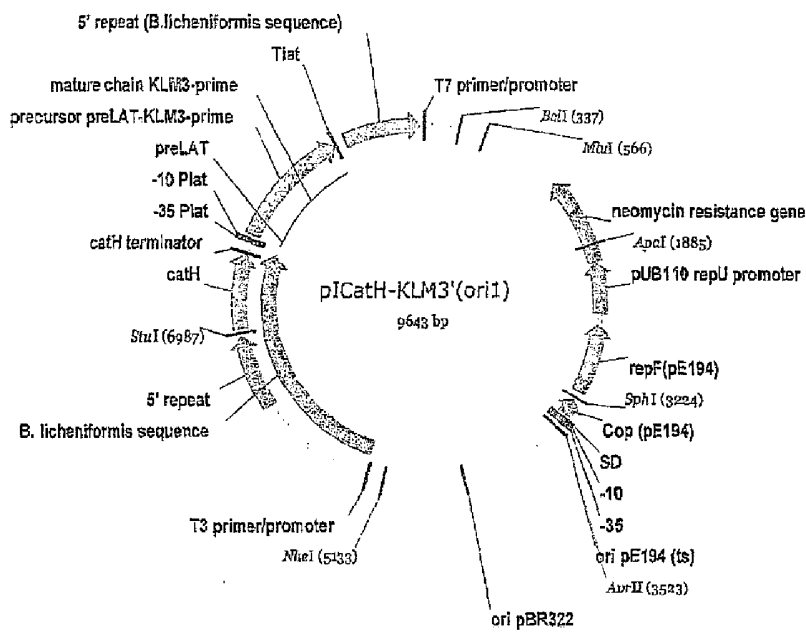


FIGURE 55

-35

```

1  GCTTTTCTTT TGAAGAAA TATAGGAAA ATGGTACTTG TTAATAATTC GGATATTTTA
   CGAAAGAAA ACCTTCTTTT ATATCCCTTT TACCATGAAC AATTTTAAAG CCTTATAART
   -10
61  TACAATATCA TATGTTTCAC ATTGAAGGG GAGGAGAATC ATGAAACAAC ABAACGGCT
   ATGTTATAGT ATACAAGTG TAACTTTCCC CTCCTCTTAG TACTTTGTTG TTTTTCGGGA
   * Y A R L L T L L F A L I F L L F R S A A
121  TTACGCCCGA TTGCTGACGC TGTATTTCG GGTCACTTC TTGCTGCTTC ATTCTGCAGC
   RATGCCGGCT AACGACTGCG ACAATAAACG CAGGTAGAAG AACGACGGAG TARGACCTCG
   * S A A D T R P A F S R I V M F G D S L S
181  TTCAGCAGCA GATACAGAC CGGCGTTTG CCGGATCGTC ATGTTTGSAG ATAGCCTGAG
   AAGTCSTCGT CTATGTTCTG GCGGCAATC GGCCTAGCAG TACAAACCTC TATCGGACTC
   * D T G K M Y S K M R G Y L P S S P P Y Y
241  CGATACGGGC AAAATGTATA GCAAATGAG AGGCTATCTT CCGTCAAGCC CGCCGTATTA
   GCTATGCCCG TTTTACATAT CTTTTTACTC TCCGATAGAA GGCAGTTCGG GCGGCATART
   * E G R F S N G P V W L E Q L T K Q F P G
301  TGAAGGCCGC TTATGCAATG GACCGGTCG GCTGGAACAA CTGACGAAAC AATTTCCGGG
   ACTTCCGGCG AATCGTTTAC CTGGCCAGAC CGACCTGTTT GACTGCTTTG TTAAGGACCC
   * L T I A N E A E G G A T A V A Y N F T S
361  ACTGACCATC GCTAATGAAG CAGAGGAGG AGCAACAGCC GTCCCTTATA ACATAATCAG
   TGACTGCTAG CGATTACTTC GTCTTCTCC TCGTTGTGCG CAGCGGATAT TGTTTTAGTC
   * W D P K Y Q V I N N L D Y E V T Q F L Q
421  CTGGGACCGC AATATCAGG TCATCAACAA CCTGGACTAT GAAGTCACAC AGTTTCTTCA
   GACCTTGGGC TTTATAGTCC AGTAGTTGTT GCACCTGATA CTTCAGTGTG TCAAGAAAGT
   * K D S F K P D D L V I L W V G A N D Y L
481  GAAAGACAGC TTTAAACCGG ATGATCTGCT CATCCTTTGG GTCCGCGCCA ATGATTATCT
   CTTTCTGTCC AATTTGGGCC TACTAGACCA GTAGGAACCC CAGCGCGGCT TACTAATAGA
   * A Y G W N T E Q D A K R V R D A I S D A
541  GGCCTATGCG TGGACACAG AACAAATGCG CAAGAAGATC AGAGATGCCA TCAGCGATGC
   CCGCATACCG ACCTTGTGTC TTGTTCTAGG GTTCTCTCAG TCTCTAGGCT AGTCGCTACG
   * A N R M V L N E A K Q I L L F N L P D L
601  CGCTAATAGA ATGGTCCCTGA ACGGCGCCAA ACAATCCTG CTGTTTAAAC TGCCGGATCT
   GCGATTATCT TACCGGACT TGCCCGCGTT TGTTTAGCAC GACAAATGG ACCGCTTAGA
   * G Q N P S A R S Q K V V E A V S H V S A
661  GGGACAAAT CCGAGCGCCA GAAGCCAAA AGTCGTGGA GCAGTCAGCC ATGTCAGCGC
   CCCTGTTTTA GGCTCGCGCT CTTCGGTTTT TCAGCAGCTT CBTAGTCCG TACAGTCCG
   * Y H N K L L L N L A R Q L A P T G M V K
721  CTATCATAC AACTGTGTC TGAACCTGGC AAGACATTC GCACCGACGG GAATGTATTA
   GATAGTATTG TTTGACGAGC ACTTGGACCG TTCTGTTAAC CCGTGGCTGCC CTTACCAATT
   * L F E I D R Q F A E M L R D P Q N F G L
781  ATGTTTGA AATTGACAAAC AGTTTGCCGA AATGCTGAGA GATCCGCAAA ATTTTGGCCT
   TAACAACCTT TAAGTGTTC TCBAACGGCT TTACGACTCT CTAGGCGTTT TAAACCGGA
   * S D V E N P C Y D G G Y V W K P F A T R
841  GAGCGATGTC GAAGACCGCT GCTATGATGC CCGATATGTC TCGAACCCTT TTGCCACAG
   CTCGCTACAG CTTTGGGCA CGATACTACC GCGTATACAG ACCTTTGGCA AACGGTGTTC
   * S V S T D R Q L S A F S P Q E R L A I A
901  AAGCGTCAGC ACGGATAGAC AACTGTCAGC GTTACCCCG CABGAAGAC TGGCAATCGC
   TTCGAGTCC TGCCATATCG TGCACATCG CAATTCGGGC GTTCTTCTG ACCGTAGCG
   * G N P L L A Q A V A S P M A R R S A S P
961  CGGAATCCG CTTTGGCAC AAGCAGTTC TTCACCGATG GCAAGAGAT CAGCAGCCCG
   GCCTTTAGGC GAAGACCGTG TTCGTCAACG AAGTGCTTAC CGTTCTTCTA GTCTTCGGG
   * L N C E G K M F W D Q V H P T T V V H A
1021  GCTGANTGCG GAAGGCAAAA TGTTTTGGCA TCGGTCCTAT CGGACACAG TTGTCCATGC
   CGACTTAACG CTTCCGTTTT ACAAAACCTT AGTCCAGGTA GGCTGTTGTC AACAGETACG
   * A L S E R A A T F I E T Q Y E F L A H G
1081  TGCCCTTCA GAAGAGCGG CGACGTTTAT CGAACACAG TATGAATTTC TGGCCCATGG
   ACGGGAAGT CTTTCTGCC GCTGCAATA GCTTGTGTC ATACTTAAG ACCGGCTACC
   *stop
1141  CTGAGTTAAC AGAGGACGGA TTCTCTGAAG GAAATCCGTT TTTTATTTT AAGCTTGGAG
   GACTCAATTG TCTCTGCTT AAGGACTTC CTTTAGGCA AAAAAAATAA TTCGAACCTC
1201  ACAAGGTAAA GGATAAAACC TCGAG
   TGTTCATTT CTTATTTTGG AGCTC

```

FIGURE 56

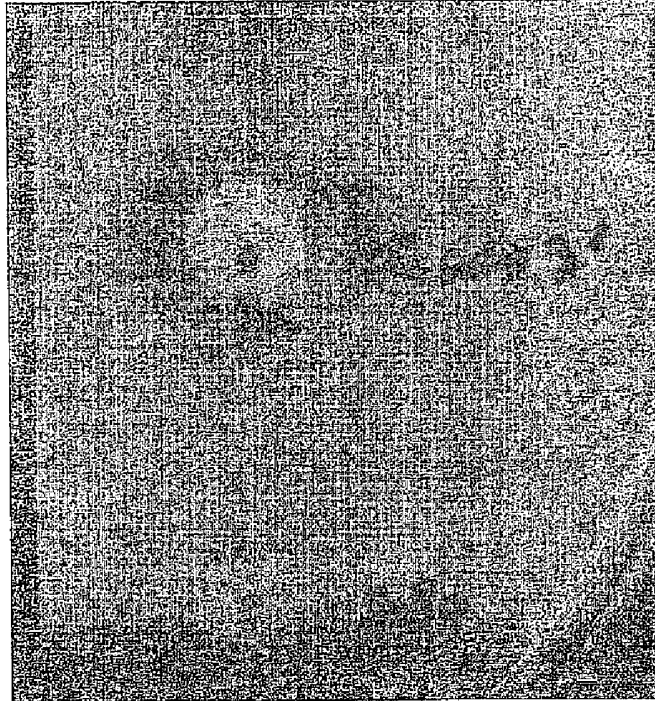


FIGURE 57 (SEQ ID No 49)

```

1   ATGAACACAC AAAAACGGCT TTACGCCCGA TTGCTGACGC TGTATTTTGC
    TACTTTGTTG TTTTTCGCGA AATGCGGGCT AACGACTGCG ACAATAAACG

51  GCTCATCTTC TTGCTGCGTC ATTCTGCAGC TTCAGCAGCA GTATCAGAC
    CGAGTAGAAG AACGACGGAG TAAGACGTCG AAGTCGTCGT CTATGTTCTG

101 CGGCGTTTAG CCGGATCGTC ATGTTTGGAG ATAGCCTGAG CGATACGGGC
    GCGGCAATC GGCCTAGCAG TACAAACCTC TATCGGACTC GCTATGCCCG

151 AAAATGTTTA GCAAAATGAG AGGCTATCTT CCGTCAGGCC CGCCGTATTA
    TTTTACATAT CGTTTTCCTC TCCGATAGAA GGCAGTTCGG GCGGCATAAAT

201 TGAAGGCCGC TTAGCAATG GACCGGTCG GCTGGAACAA CTGACGAAAC
    ACTTCCGGCG AAATCGTTAC CTGGCCAGAC CGACCTTGTT GACTGCTTTG

251 AATTTCCGGG ACTGACGATC GCTAATGAAG CAGAAGGAGG AGCAACAGCG
    TTAAAGGCCG TGACTGCTAG CGATTACTTC CTCCTCCTCC TCGTGTCTCG

301 GTGCGCTATA ACAAAATCAG CTGGGACCGG AAATATCAGG TCATCAACAA
    CAGCGGAZAT TGTTTTAGTC GACCTTGGGC TTTATAGTCC AGTAGTTGTT

351 CCTGGACTAT GAAGTCACAC AGTTTCTTCA GAAGACAGC TTAAACCGCG
    GCACCTGATA CTTCAGTGTG TCAAAGAAAT CTTTCTCTCG AAATTGSCC

401 ATGATCTGGT CATCCTTGG GTGCGCGCCA ATGATTATCT GCGTATGGC
    TACTAGACCA GTAGAAACCC CAGCGCGGGT TACTAATAGA CCGCATACCG

451 TGAACACAG AACAAATGC CAAAAGATC AGAGATSDCA TCAGCGATGC
    ACCTTGTGTC TTGTCTACG GTTTTCTCAG TCTCTACGGT AGTCGCTACG

501 CGCTAATAGA ATGCTCTGA ACGGCGCCAA ACAATCTCTG CTGTTTAAAC
    CCGATTATCT TACCAGGACT TCGCGCGGTT TGTTTAGGAC GACAATTTGG

551 TGCGGGTCT GGGACAAAT CCGAGCGCCA GAGCCGAAA ACTCGTCGAA
    ACGGCCTAGA CCTGTTTTA GGCCTCGGCT CTTGCTTTT TCAGCAGCTT

601 GCAGTCAGCC ATGTCAGCCG CTATCATAAC AAATGCTGTC TGRACCTGGC
    CGTCAGTCGG TACAGTCGG GATAGTATTG TTTGACGACG ACTTGGACCG

651 AAGACAATTG GCACCGACGG GAATGGTTAA ATTGTTTGA ATTGACAAAC
    TTCTGTTAAC CBTGGCTGCC CTTACCAATT TAAACAACTT TACTGTTTG

701 AGTTTGGCGA AATGCTGAGA GATCGGCAA ATTTTGGCCT GAGCGATGTC
    TCAPACGGCT TTACGACTCT CTAGCCCTTT TAAACCGGA CTCGCTACAG

751 GAAACCCGT GCTATGATGG CCGATATGTC TGSAAACCGT TTGCCACAAG
    CTTTGGGCA CGATACTACC GCCTATACAG ACCTTTSGCA AACGCTGTTG

801 AAGCGTCAGC ACGGATAGAC AACTGTCAGC GTTTGGCCCG CAAGAAAGAC
    TTGCGAGTCG TGCCATATCG TTGACAGTCG CAATCGGGC GTTCTTTCTG

851 TGGCAATCGC CGAAATCCG CTTTGGGCA AAGCAGTTGC TTCACCGATG
    ACCGTTAGCG GCGTTTAGGC GAAACCGTG TTCGICACG AAGTGGCTAC

901 GCAAGAGAT CAGCAAGCCC GCTGAATTGC GAAGCAAAA TGTTTTGGGA
    CGTTCTTCTA GTCGTTGGGG CGACTTAACG CTTCCGTTTT ACAAAACCTT

951 TCAGGTCCAT CCGACACAG TTGTCCATGC TGCCCTTTCA GAAGAGCGCG
    AGTCCAGGTA GGCTGTTGTC AACAGGTACG ACGGGAAAGT CTTTCTCGCC

1001 CGACGTTTAT CGAAACACAG TATGAATTC TGGCCCATGG CTGA
    GCTGCAATA GCTTGTGTC ATACTTAAG ACCGGGTACC GACT

```

FIGURE 58 (SEQ ID No. 50)

```

1  ATGAAAAAT  GCTTTGTGTC  TTTATTGGGA  TTGCTCGCGC  TGACAGTTCA  GGCAGCCGAC
61  AGCCGTCCCG  CCTTCTCCCG  GATCGTGATG  TTTGCGGACA  GCGTCTCCGA  TACCGGCAAG
121  ATGTACAGCA  AGATGCGCGG  TTACCTCCCG  TCCAGCCCGC  CCTACTATGA  GGGCCGCTTC
191  TCCACGCGCG  CCGTCTGGCT  GGAGCAAGCTG  ACCAAGCAGT  TCCCGGCGCT  GACCATAGCC
241  AACGAGGCGG  AAGGCGGACC  GACCGCCGTG  GCTTACAACA  AGATCTCCCTG  GATCCCAAG
301  TATCAGGTCA  TCACACACCT  GGACTACGAG  GTCACCCAGT  TCGTGCAAAA  AGACAGCTTC
361  AAGCCGGACG  ATCTGGTGAT  CCTCTGGGTC  GGCGCCAAG  ACTATCTGGC  CTATGGGTGG
421  AACACAGAGC  AGGATGCCAA  GCGGCTGCGC  GACGCCATCA  GCGATGCGGC  CAACCGCATG
481  GTGCTGAAGC  GCGCCAAGGA  GATACTGCTG  TTCACCTGCG  CGGATCTGGG  CCAGAACCCG
541  TCGGCGCGCA  GCGAGAAAGT  GGTGAGGGCG  GCGAGCCATG  TCTCCGCGTA  CCACAAACAG
601  CTGCTGCTGA  ACCTGGCAGC  CCAGCTGGCT  CCCACCGGCA  TGGTGAAGCT  GTTCGAGATC
661  GACAAGCAGT  TTGCGCAGAT  GCTGCGTGAT  CCGCAGAACT  TCGGCGTGAG  CGACCAAGAG
721  AACGCTGCT  ACGGTGGCAG  CTATGTGTGG  AAGCCGTTTG  CCTCCCGCGG  CGCCAGCACC
781  GACAGCCAGC  TCTCCGCCCT  CAACCCGCGG  GAGCGCCTCG  CCGTCCGCGG  CAACCCGCTG
841  CTGCGCCAGG  CCGTCCGCCG  CCCATGGGCT  GCGCGCAGCG  CCGACCCCTT  CAATCTGAGG
901  GGGAGATGT  TCTGGGATCA  GGTCCACCGC  ACCACTGTGG  TGCACGCCCG  CCTGAGCGGG
961  CCGCGCGGCA  CTTTCATCGA  GAGCCAGTAC  GAGTTCTCTG  CCCAC

```

FIGURE 59 (SEQ ID No. 51)

```

1  ATGAAAAAT  GCTTTGTTTG  TTTATTGGGG  TTGATCGCGC  TGACAGTTCA  GGCAGCCGAC
61  ACTCGCCCGG  CCTTCTCCCG  GATCGTGATG  TTGCGCGACA  GCGTCTCCGA  TACCGGCAAG
121  ATGTACAGCA  AGATGCGCGG  TTACCTCCCG  TCCAGCCCGC  CCTACTATGA  GGGCCGCTTC
181  TCCACCGGAC  CCGTCTGGCT  GGAGCAAGCTG  ACCAAGCAGT  TCCCGGCTCT  GACCATCGCC
241  AACCAAGCGG  AAGGCGGTGG  CACTGCGGTC  GCTTACAACA  AGATCTCCCTG  GATCCCAAG
301  TATCAGGTCT  ACACACACCT  GGACTACGAG  GTCACCCAGT  TCTTGCAAAA  AGACAGCTTC
361  AAGCCGGACG  ATCTGGTGAT  CCTCTGGGTC  GGTGCCAATG  ACTATCTGGC  ATATGGGTGG
421  AATACCGGAG  AGGATGCCAA  GCGGCTGCGC  GATGCCATCA  GCGATGCGGC  CAACCGCATG
481  GTACTGAAGC  GTGCCAAGCA  GATACTGCTG  TTCACCTGCG  CGGATCTGGG  CCAGAACCCG
541  TCAGCCCGCA  CTCAGAAAGT  GGTGAGGGCG  GTCAGCCATG  TCTCCGCGTA  TCACAAACAG
501  CTGCTGCTGA  ACCTGGCAGC  CCAGCTGGCT  CCCACCGGCA  TGGTGAAGCT  GTTCGAGATC
561  GACAAGCAAT  TTGCGCAGAT  GCTGCGTGAT  CCGCAGAACT  TCGGCGTGAG  CGACCTGAGG
621  AACCCCTGCT  ACGAGCGCGG  CTATGTGTGG  AAGCCGTTTG  CCGCCCGCAG  CGTCAACACC
681  GACCGCCAGC  TCTCCGCCCT  CAGTCCGCGG  GAAAGCCTCG  CCGTCCGCGG  CAACCCGCTG
741  CTGCGACAGG  CCGTTGCCAG  TCCATATGGC  CCGCGCAGCG  CCAGCCCGCT  CAATCTGAGG
801  GGCAGAGAT  TCTGGGATCA  GGTACACCGC  ACCACTGTGG  TGCACGCCCG  CCTGAGCGGG
861  CCGCGCGGCA  CTTTCATCGA  GAGCCAGTAC  GAGTTCTCTG  CCCACGAGG  A

```

FIGURE 60 (SEQ ID No. 52)

```

1  ATGCGGAGAG  CTGCGCTTGG  CCGTGCTATG  ACCGCGAGAG  TCGCGCGCGT  CGGCGCGCTC
61  GCGCTCGGCG  TCACCGAGCG  CACCGGCCAC  GCGCGCGCGG  CCGAGGCGAC  TCGGACCTTG
121  GACTACGCTG  CCGTCCGCGA  CAGCTACAGC  GCGCGCTCCG  GCGTCTCTGC  CGTCCGACCC
181  GCGCAACCTG  TCTGTCTGGG  CTCGACGGCG  AACTACCGCG  ACCTCATCGC  GGCACAGAGG
241  GCGCGCGCGC  TCACGGAGCT  CAGCTGCGGC  GCGCGCGAGA  CCGCGGACTT  CAGCGCGCGC
301  CAGTACCGCG  GCGTGGCAGC  CCGTGTGGAG  GCGCTCGGCA  CCGGCAAGGA  CCGTGTGAGG
361  CTCACCATCG  GCGGCAACGA  CACAGGACCG  TTCATCAACG  CCGTACCGCG  CTCGCGCAGG
421  GCGGCTGTCC  TCAGCGCGCG  CAGGCGGAGC  CCGTGCAGAG  ACAGGCGAGG  CACCTCTCTC
481  GACGACGAGA  TCGAGGCGAA  CAGCTACCGC  GCGCTCAAGG  AGGCGCTGCT  CGCGCTCGCG
541  GCGAGGCGTC  CCGACGCGAG  GGTGGCGGCT  CTGCGCTACG  CGTGGATCAC  CCGGCGCAGC
601  GCGGCGCGCT  CCGTCTCTCT  GAGCTTCGCG  CTCGCGCGCG  GTCAGCTGCC  CTACCTGCGG
661  GCGATCCAGG  CACACCTCAA  CAGCGCGGTC  CCGCGGCGCG  CCGAGGAGAG  CCGAGCGCAC
721  TACGTTGAGT  TCTCCGCGGT  GTCCGACGGC  CAGGACGCGT  CCGAGGCGCG  CCGGACCGCG
781  TGGATCGAAC  CGCTGCTCTT  GCGGCGACAG  CTGCTTCGCG  TCCACCCCAA  CGCGCTGCGG
841  GAGCGCGGCA  TGGCGGAGCA  CAGGATGAGC  GTCTCGGCGC  TGGACTGA

```

FIGURE 61 (SEQ ID No. 53)

```

1   TCAGTCCAGG CCGAGGACGT CCACTCGTGT CTGGGCCATG CGCCGCTCGC CCAGGGCGTT
61  GGGGTGGACG GGAACGAGGC TGTGCCCGAA GAGCAGGGT TCGATCCAGC GGGTGCCTGG
121 GGCCTCGCAG GCGTCTTGGC CGTGGGACAC CCGGAGAGG TCCACGTAGG TGGCTCCGGT
181 CTCTCTGGCG GCGCGCGGGA CCGCGTCTGT GAGGTGTGCG TGGATGGCCC GCAGGTAGGG
241 CRCTCACCG GCGGCGAGGG GGAGCTTCAG GAGCAGGAC GGGTCCGGGG TGGCCGGGGT
301 GATCCACGGG TAGCGGAGAG CCGCCACCGT GCGCTGSSGA GCGCTGGGCG GGAAGCCGAG
361 CAGCGCTTCC TTGAGCGCGG GGTACGTGTT GGCCTCGATC TCGTCTGTGA AGGAGGTGGC
421 GTGCTTSTCC TTGACGGGGC TGCCTTTCGC GCGCTGAGG ACACCCGCGC TSCCGCAGGC
481 CGTGATGGCG TTGATGAAGG TGCTGTGTGC GTTGCCGCGC ATGCTGAGCG TGACCAGGTC
541 CGTSCCGGTG CCGAGCGCGT CCAACTCGGG TSCGACGCGC GGGTACTGGG CCGCGGTGAA
601 GTCCGCGGTC TSCCGCGCGC CGCAGGTGAC GTCCGTGAGG CCGGCGCGCG TCGTGTCCGC
661 GATGACGTGG GGTAGTTTGG CCGTCGAGCG CAGACGAGCG AGGTGCGCGG GGTGACGGGG
721 CAGGACGCGG GAGCGGGCGC TGTAGCTGTC GCGGAGGGCG ACCTAGTCCA GGGTCGGAGT
781 GGCCTGGGCG GGCSCGCGCT CCGCGGTGGC GTCGGTGAGG CCGAGGGGGA GCGTGCCTGC
841 GCGGCGGACT TCGCGGTCA TGACACGGCG AAGGGCAGGC TTCGGCAT

```

FIGURE 62 (SEQ ID No. 54)

```

1   ATGGATTACG AGAAGTTTCT GTTATTGGG GATTCCATTA CTGARTTTCG TTTTAATACT
61  AGGCCCAATG AAGATGGCAA AGATCAGTAT GCTCTTGAG CCGCATTAGT CAACGAATAI
121 ACGAGAAAAA TGGATATCTT TCAAAGAGGG TTCAAAGGGT ACACTTCTAG ATGGGCGTTG
181 AATAACTCTC CTGAGATTTT ANAGCATGAA TCCAAATTG TCATGGCCAC AATATTTTGG
241 GGTGCCAACG ATGCATGCTC AGCAGGTCCG CAAAGTGTCC CCGTCCCGGA ATTTATCGAT
301 AATATTCTGC AATGGTATC TTGATGAGG TCTTACCATA TCCGTCTAT TATAATAGGA
361 CCGGGGCTAG TAGATAGAGA GAAGTGGGAA AAGAAATAAT CTGAGAAAT AGCTCTCGGA
421 TACTTUCCTA CCAACGAGAA CTTPGCCAT TATTCGGATG CCTAGCAAA ACTAGCCAAI
481 GAGGAAAAAG TTCCCTTCGT GCGTTTGAT AGGCGTTTC AACAGGAGG TGGGTATGCT
541 TSGGACAAAC TGCTAACAGA TGGACTGCAC TTTPCCGGA AAGGATCAA AATTTTCAT
601 GACGAATTAT TGAAGGTCAI TGAGACATTC TACCCCAAT ATCATCCCAA AATCATGCAG
661 TACEAACTGA AAGATTGCGG AGATGTGCTA GATGATGGAT CTAACATAT GTCTTGA

```

FIGURE 63 (SEQ ID No. 55)

```

atgaacctgc gtcgaatggat gggcgccgcc acggctgccc ttgacctggg cttggccggc      60
tgccggggcg gtgggacgga ccagagcggc aatcccaatg tcgccaaggt gcagcgcatg      120
gtggtgttcg gcgacagcct gagcgatata ggcacctaca ccccgctcgc gcaggcggtg      180
ggcgccggca agttcaccac caaccgggca ccgatctggg ccgagacggt ggcggcgcaa      240
ctggcggtga gcgtcacgcc ggcgggtgat gctcacgcca cctccgtgca gaattgcccc      300
aagggccggct gcttcgacta tgcgcagggc ggctcgcgcg tgacgatcc gaacggcatc      360
ggcacaacag gcggcgcggg ggcgctgacc taccgggttc agcagcgct gcgcaacttc      420
tacggcgcca gcaacaacac attcaacgga aataacgatg tcytcttcgt gctggccggc      480
agcaacgaca tttctctctg gaccactgag gggccacca cgggctccgg cgtgacgccc      540
gccattgcca cggcccaggt gcagcaggcc gcgacggacc tggctcgcta tgtcaaggac      600
atgatcgcca aggggtcgac gcaggtctac gtgttcaacc tgcgcgacag cagcctgacg      660
ccgagcggcg tggcaagcgg cagcagcgcc caggcgctgc tgcacgcgct ggtgggacag      720
ttcaacacga cgctgcaaa ggggctggcc ggcacgtcgg cgcgcacat cgaactcaac      780
gcacaactga ccgcggcgat ccagaatggc gctcgttcg gcttcgcaa caccagcgcc      840
cgggctcgcg acgccaccaa gatcaatgcc ctggtgcgca gcgcggcgcg cagctcgctg      900
ttctgctcgg ccaacacgct ggtggcttcc ggtgcggacc agagctacct gttcgccgac      960
ggcgtgcaco cgaccacggc cggccatcgc ctgatcgcca gcaacgtgct ggcgcgctg      1020
ctggcgata acgtcgcta ctga      1044

```

FIGURE 64 (SEQ ID No. 56)

```

1 gtgatcgggt cgtacgtggo ggtgggggac agcttcaccc agggcgctcgg cgaacccggc
61 cccgacgggg cgttcgtcgg ctgggcccag cggctcgcgg tactcgtcgg ggacgggcgg
121 cccgaggggg acttcacgta caggaacctc gccgtgcggg gcaggctcct cgaaccagtc
181 gtggcggaac aggtcccgcg cgtcctcgga ctccgcgcgg acctcgtctc gttcggggcg
241 gggcgcaacg acatcatccg gcccgccacc gatcccgacg aggtcgccga cgggttcgag
301 ctggcggtgg cggcgctgac cggcggggcc ggaaccgtcc tggtagccac cgggttcgac
361 acccgggggg tgcccgctct caagcacctg cggcgcaaga tggccacgta caacggggac
421 gtccgcggca tcgcccagcc ctacggctgc cgggtgctcg acctggtgtc gctcgggagc
481 gtccaggacc caggggcgtg ggaagccgac cggctgcacc tgtcccgga ggggcatacc
541 cgggtggcgc tgcgcgcggg gcaggccctg ggcctgcggc tcccggccga cctgaccag
601 cccggccgcg ccccgccgac gcggcgacg ctccagcttc ggcgggacga cgtgcactgg
661 gcgcgcgagt acctggtgac gtggatcggg cggcggtgcg gggcgagatc gtcggcgacg
721 cactgacgg ccaaggggac gctgtgcgg gacccatca agacgcggat cgcgcgggtg
781 gctga

```

FIGURE 65 (SEQ ID No. 57)

```

1 atgcagacga accccggcta caccagtctc gtgcgcgtcg gcgactcctt caccgagggc
61 atgtcggaac tctgcccga cggctcctac cgtggctggg ccgacctcct cggcaccggg
121 atggcgggcc gctcccccgg ctcccggtac gccaacctgg cggtagcggg gaagctgac
181 ggacagatcg tcgacgagca ggtggagctg gccgcggcca tgggagccga cgtgatcacg
241 ctggtcggcg ggctcaacga caggtcgcg cccaagtgcg acatggcccg ggtcggggac
301 ctgctgaccc aggcgtgga cgggtctggc ccgcactgcg agcagctggt gctgatgcgc
361 agtcccggtc gccagggtcc ggtgctggag cgtctccggc cccgcatgga ggcctgttc
421 gccgtgatcg acgacctggc cgggcggcac ggcgcgtgg tegtcgacct gtacggggcc
481 cagtgcgtgg ccgacctcg gatgtgggac gtagaccggc tgcaactgac cgcggagggc
541 caccggcggg tcggggaggg ggtgtggcag tegctcggcc acgagcccca ggaaccggag
601 tggcacgcgc cgtaccggc gacgcccggc cgggggtggg tgacgcgcag gaccggggac
661 gtccgggttg cccggcagca cctgctgccc tggataggcc gtaggtgac cgggcgctcg
721 tccggggacg gcttcccggc caagcgcccg gacctgtgc cctacgagga ccccgcacgg
781 tga

```

FIGURE 66 (SEQ ID No. 58)

```

1 atgacccggg gtcgtgacgg gggtagggg gcgcscacca caaagcacgg tgccctgctc
61 gggcgatcg tcacctgat agtggcgatc tcggcgccca tatacgccgg agcgtccggg
121 gacgacggca gcagggacca cgcctgcag gccggaggcc gtctcccacg aggagacggc
181 gcccccggct ccacgggtgc ctgggtgggc gcctgggcca ccgacccggc cgcggccgag
241 cggggcaccc agacgacgg cctggcgggc cgttcctgac gcaacgtcgt gtaacctcg
301 gtccggcgga cggcgcgcg gatcacctc tcgaacctgt accggcgatc gtcgctgacc
361 gtacacacag cctcgatcgc cctggccgac gggcccgaca cgcgcgcgcg gatcggcgac
421 accatgcgac ggctcacctt cggggcgagc gcccggtgta tcctcccgcc gggcgccag
481 gtgatgagcg acaccgccc cctcgccatc cctacgggg cgacgtcctt ggtacacacg
541 tactccccc tcccgctcgg gccggtgacc tactatccgc agggccggca gaccagctac
601 ctggcgacg gcgacggcac ggcggacgtc accgccttuy cgtacacacac cccacggccc
661 tactggcgt acctgacggc cctcgacgtg ctgagccacg agggcgacgg caccgtcgtg
721 ggttcggcg actccatcac cgaagcgccc cgttcgcaga ggcagccaa ccaacctggt
781 accgacgtcc tcggcgacg cctgcacgag gggcgggggc accggcgga caccggccgc
841 tacagcgtcg tcaacgaggg catcacgggc aaccggctcc tgaccaagag gccggggcg
901 ccggcgaca acccgagcgg actgagccgg ttccagcggg acgtgctgga acgacacac
961 gtcaaggccc tctgtcgtgt cctcgccctc aacgacgtcc tgacagccc ggaactcgcc
1021 gaccggcagc ccatctgac cggcctgcgc accctcgtcg accgggggca cggccgggga
1081 ntgggggtcg tcggcgccac gatcacggc ttccggcggt accggcgcta caccgaggtc
1141 cggagacga tggcgccaga ggtcaacgag gagatccgct ccggcggggt ctccgacag
1201 gtctcgaact tcgacaaggg cctgcgcgac ccgtacgacc cgcgcgggat ggcctccgac
1261 tacgacagcg ggaacccct gcaccccgcc gacaagggtt agcgcgcat gggcgcggtc
1321 atcgacctgg ccggcgtgaa gggcggggg cgggtcaagg cgtag

```

FIGURE 67 (SEQ ID No. 59)

```

1  atgacagagca tctcagagggc gagggtggcg cggcggatcg cggccggcgc ggctacggc
61  ggccggcgac tccggctggc gggagcggcg gcggtcggtc tgggtggcg cgaggtgacg
121  ctggccagac gcaggtggg ggtgggcacg ccgacccggg tgcggaacgc gcagggactg
181  tacggcgagca cctgcccac ggcggcgac cgcgcgctgc ggctgatgat gctggcgac
241  tccagggcgc cggggcagg cgtccacgg gcccggcaga cgcggcgcg gctgctggcg
301  tccgggctcg cggcggtggc ggaaggcgcg gtgcggctgg ggtcggtcgc ccagccgggg
361  gctgtgctcg acgacctgga cggcgaggcg gcgctggtgc tgcggcgcc gcacccgggtg
421  cccgacatct gctgcatcat ggtcggcgcc aacgacgtca cccacccgat gccggcgacc
481  cgctcggtgc ggcacgtgc ctggcggtg cggcggtgc gcacggccgg tccggagggtg
541  gtggctggca cctgtcggg cctgggcacg atcgagcggg tgcggcagcc gctgctggtg
561  ctggccggcg gggcctcacg gcagctcgcg ggggcacaga ccacggcgcg cgtcgagcag
661  ggcggcgagc cgggtcgctt gggcgacgtg ctgggtccgg agttcgcgca gaacccggcg
721  gagctcttcg gcccgcacaa ctaccacccc tccggcgagg ggtacggccc gccggcgatg
781  ggggtactgc cctcggtgtg cgcgcgcctc ggcctgtggc cggcgagcga ggcgcacccg
841  gacgcgctgc gccgcgaggg ctctctgcgg gtggcgcgcg cggcgcgaga gccggcgctc
901  gaggcggtga cggaggtcgc cgcgcacatg cctacggggc ctccggcgcc ctggcgctg
961  ctgaagcgcc ggaagcggcg tccgggtgtg gaggcggaac cgtccagccc gtcggcgctt
1021 tga

```

FIGURE 68 (SEQ ID No. 60)

```

1  atgggtcgag ggacggacca ggggacggcg tacggcgctc gccggcgcg tgtcgcgctc
61  gccgcctga cgcgcgcgt cctggcggtg ggcgtggcg cctgcgactc cgtggcgcg
121  gactcaccgc ctctctccg cagcccgctg aagcggacga ggcggcgcc cgcctgggac
181  accagcccg cgtccgtgc cgcgtggcg gactccatca cgcggcgctt cgaagcctgt
241  gcggtgctgt ccgactgccc ggggtgtcg tggcgacgg gccgcagcg gaaggtcgac
301  tgcgtggcg tanggtgct ggggaaggcg gacggcgcc agcagcgtg gaactacgg
361  gtcacgggg cccggatggc ggaactgacc gctcagggtg cgcggcgcg gcagcgcgag
421  ccggagctgg tggcggtgat ggcggggcg aacgacgctg gccggtccac gaactcggcg
481  atgacggcg tggcgactt cggcgcgag ttcgaggagg cgtggccac cctgcgcaag
541  aagctcccca aggcgcaggt gtacgtgtcg agcatcccg acctcaagcg gctctggtcc
561  caggggcgca ccaacccgct gggcaagcag ggtggaagc tccgctgtg cccgtcgatg
661  ctggcgagcg cggactccct ggaactcgcg gcgaactgc ggcgcacac ggtgcgagc
721  cgggtggcg actacaacga ggtgctcgcg gaggctcgc cgaaggacgg gcgtgccc
781  agcgacgag cgcgggtgca cgaattccg ttcggcacgg accagttgag ccatgggac
841  tggttccacc cgaattgga cggccaggcc cggctggcg agatcgcta cgcgcgggtc
901  accgcgaaga atccctga

```

FIGURE 69 (SEQ ID No. 61)

```

1  ttcatcaca cgaatgcaca acacgggcca tccgggicac cctgatcgt ggggaatggg
61  gacagacgtt cccgtgacga aagggtcctg ctacatcaga aatgacagaa atcctgctca
121  gggaggttcc atgagactgt cccgacgcgc ggcacggcg tccgcgctcc tccctacccc
181  ggcgctcgcg ctcttcggcg cgaagcggcg cgtgtcgcg cgcgaatcc aggcaccca
241  ctactggcg ctcggcgact cctactctc ggggtcggc ggggscagct acgacagcag
301  cagtggctcc tgaagcgca gcaccaagtc ctacccggcc ctgtggggcg cctcgccac
361  cggtagcgcg ttcaacttca ccgctgttgc gggcgccgc acaggagagc gctgtggcaa
421  gcagctgacc ccgtcaact cgggcacgga cctggtcagc attaccatcg gcgcgaacga
481  cgggggttcc gccgaccca tgaccacgtg caacctccag ggcgagagcg cgtgcttggc
541  gcgatcgcc aaggcgcgcg cctacatcca gcagacgctg cccgcccagc tggaccaggt
601  ctacgacgac atcgacagcc gggccccgc agcccaggtc gtcgtcctgg gctacccgg
661  cttctacaag ctggcgcgga gctgcgcgt cggctctcgc gagaagtccc gcgcggccat
721  caacgcgcgc gccgacgaca tcaagcgcgt eacgcgcaag cgcgcgcgg accacggctt
781  cgccttcggg gacgtcaaca cgaacttgc cgggcacgag ctgtgctcgg gcgcgcctg
841  gctgcacagc gtcacccctc cgtggagaa ctctaccac cccacggcca acggacagtc
901  caagggctac ctgcccgtcc tgaactcgc caactgatct cgcggctact cgcgccttga
961  cgaagtccg ccccgggcg gggcttcgac gtaggtgccc gtacggcggt cgcgcgtcgc
1021 gccgggtggc cgcgctacg tgcgcgcgc cccggagcgc gtcggttc

```

FIGURE 70 (SEQ ID No. 62)

```

1  ATGAAAAAAT  GGTTCGTGTG  TTTATTTGGA  TTGGTCCGCG  TGACASTTCA
   TACTTTTTTA  CCAACACAC  AATTAACCC  AACACGCGCG  ACTGTCAAGT

51  GGCAGCGGAC  AGTCGCCCG  CTTTTCCTCG  GATCGTGTAT  TTCGGCGACA
   CCGTCGCGTG  TCAGCGCGG  GGAAGAAGGC  CTAGCACTAC  AAGCCGCTGT

101  GCCTCTCCGA  TACCGGCAAA  ATGTACAGCA  AGATCCGCGG  TTACCTCCCG
   CGGAGAGGCT  ATGCGCGTTT  TACATGTCTG  TCTACGCGCC  AATGGAGGGG

151  TCCAGCCCGC  CCTACTATGA  GGGCCGTTC  TCCAACGGAC  CCGTCTGGCT
   AGGTCCGCGG  GGATGATACT  CCCGGCAGAG  AGGTTCCTTG  GGCAGACCGA

201  GGAGCAGCTG  ACCAAACAGT  TCCCGGGTCT  GACCAATCGC  AACGAAGCGG
   CCTCGTCGAC  TGCTTTGTCA  AGGGCCAGAG  CTGCTAGCGG  TTGCTTCGCC

251  AAGGCGGTGC  CACTGCCGTG  GCTTACAACA  AGATCTCTCT  CAATCCCAGG
   TTCCGCCACG  GTGACGGCAC  CGAATGTGT  TCTAGAGGAC  CTTAGGGTTC

301  TATCAGSTCA  TCAACACCT  GGACTACGAG  GTACCCAGT  TCTTGCAGAA
   ATAGTCCAGT  AGTTGTGTGA  CCGATGCTC  CAGTGGGTCA  AGAAGCTCTT

351  AGACAGCTTC  AAGCCGGACG  ATCTGGTGAT  CCTCTGGGTC  GGTGCCATG
   TCTGTGAGAG  TTCGGCCTGC  TAGACCACTA  GGAGACCCAG  CCACGGTTAC

401  ACTATCTGGC  CTATGGCTGG  AACACGGAGC  AGGATGCCAA  GCGGGTTCGC
   TGATAGACCG  GATACCGACC  TTGTGCTCTG  TCCTACGGTT  CGCCCAAGCG

451  GATGCCATCA  GCAATGGGCG  CAACCGCATG  GTACTGAAAG  GTGCCAAGCA
   CTACGGTAGT  CGCTACCGCG  GTTGGCGTAC  CATGACTTGC  CACGGTTCTG

501  GATACTGTCT  TTCAACCTGC  CGGATCTGGG  CCAGAACCCG  TCAGCTCGCA
   CTATGACGAC  AAGTTGGACG  GCTAGACCC  GGTCTTGGGC  AGTCGAGCGT

551  GTCAGAGGTT  GGTGGAGCGG  GTCAGCCATG  TCTCCGCTTA  TCACAACCGG
   CAGTCTTCCA  CCAGCTCCGC  CAGTCCGTAC  AGAGGCGGAT  AGTGTGTGTC

601  CTGCTGCTGA  ACCTGSCACG  CAGCTGGGCG  CCCACCGGCA  TGSTAAAGCT
   GACGACGACT  TGAACCGTGC  GGTGACCGCG  GGTGGCCGCT  ACCATTTCGA

651  GTTCGAGATC  GACAAAGCAAT  TTGCCGAGAT  GCTGCGTGAT  CCGCAGAACT
   CAAGCTCTAG  CTGTTCTGTA  AACGGCTCTA  CGACSCACTA  GCGCTCTTGA

701  TCGGCGCTGAG  CGAGGTGAG  AACCCCTGCT  ACSACGGCGG  CTATGTGTSS
   AGCCGGACTC  GGTGCAGCTC  TTGGGGACGA  TGCTGCCGCC  GATACACACC

751  AAGCCCTTTC  CCACCCGCG  CGTCAGCACC  GACCCCGAGC  TCTCCGCTTT
   TTCGGCAAC  GGTGGCGTTC  GCACTCGTGG  CTGGCGCTCG  AGAGGCGGAA

801  CAGTCCGCG  GAACGCTTC  CCATCGCGCG  CARCCCGCTG  CTGGCACAGG
   GTCAGCGCTC  CTGCGCGAG  GCTAGCGGCG  GTTGGCGGAC  GACCGTGTCC

851  CCGTTGCCAG  TCTATGGCC  CGCGCGAGCG  CCAGCCCGCT  CAACTGTGAG
   GSCAACGGTC  AGGATACCG  GCGCGCTCG  GGTGCGGGGA  GTTGACACTC

901  GGCAGATGT  TCTGGGATCA  GGTACACCG  ACCACTGTCT  TGCACGCGAC
   CCGTTCTACA  AGACCTTAGT  CCATGTGGGC  TGGTGACAGC  ACCTGCCCTG

951  CCTGAGCGAG  CGCGCGCGCA  CTTTCACTGC  GACCACTAC  GAGTTCCCTG
   GCACTCGCTC  GCGCGCGGCT  GGAAGTAGCG  CTTGGTCTAG  CTCAGGAGCG

1001  CCCAC  TGA
      GGGTG  ACT

```

FIGURE 71 (SEQ ID No. 63)

```

1  ATGAAAAAT  GGTTCGTTG  TTTATTGGG  TTGATCGCG  TGCAGTTCA
   TACTTTTAA  CCAACCAAC  AATAACCC  AACTAGCGC  ACTGTCAAT

51  GGCAGCGGAC  ACTGCCCCG  CCTTCTCCG  GATCGTGAT  TTGGGGACA
   CGCTCGGCTG  TGAGCGGGG  GGAAGAGGG  CTAGCACTAC  AAGCCGCTGT

101  GCCTCTCCGA  TACCGGAAA  ATGTACAGC  AGATGCGCG  TTACCTCCG
   CGSAGAGGCT  APGGCCGTT  TACATGTCT  TCTACCGCC  AATGAGGGG

151  TCCAGCCCC  CCTACTATG  GGGCCGTTT  TCCACGGAC  CGCTCTGGT
   AGCTCGGCG  GGTGATACT  CCCGGCAAAG  AGGTTCGCT  GGCAGCCGA

201  GGAGCAGCTG  ACCAAGCAG  TCCGGGCTT  GACCATCGC  AACGAAGCG
   CCTCGTCGAC  TGGTTCGT  AGGCCCCAG  CTGGTAGCG  TTGCTTCGC

251  AAGCGGGTGC  CACTGCGTG  GCTTACAAC  AGATCTCCT  GAATCCCAAG
   TTCCGCCACG  GTAGCGCAC  CGAATGTTG  TCTAGAGGAC  CTAGGGTTC

301  TATCAGGTCA  TCAACACCT  GGACTACGAG  GTCACCCAG  TCTTGCAGAA
   ATAGTCCACT  AGTTGTTGA  CCTGATGCT  CAGTGGTCA  AGAACGTTCT

351  AGACAGCTTC  AAGCGGACG  ATCTGCTGAT  CCTCTGGGT  GGTGCCAATG
   TCTGTGGAAG  TTGGCCCTC  TAGACCACTA  GGAGACCCAG  CCACGCTTAC

401  ACTATCTGG  ATATGGCTG  AATACGGAG  AGGATGCCA  GCGAGTTCC
   TGATAGACCG  TATACCGAC  TTATGCCTC  TCCTACGGT  CGCTCAAGCG

451  GATGCCATCA  CGGATGCGC  CAACCGCAT  GTACTGAAC  GTGCCAAGCA
   CTACGGTAGT  CGCTACGCC  GTTGGCGTAC  CATGACTTC  CACGGTCTGT

501  GATACTGCT  TTCAACCTG  CGGATCTGG  CCGAACCAG  TCAGCCCGCA
   CTATGACGAC  AAGTTGGAG  GCCTAGACC  GGTCTTGGG  AGTCGGGGCT

551  GTCAGAGGT  GGTGAGGCG  GTCAGCCAT  TCTCCGCTA  TCACAACAG
   CAGTCTTCA  CCAGCTCCG  CAGTCGTAC  AGAGGCGGAT  AGTGTGTTTC

601  CTCTGCTGA  ACCTGGCAG  CCAGCTGGC  CCCACCGCA  TGGTAAAGCT
   GACGACGACT  TGGACCGTC  GGTGACCCG  GGGTGGCCGT  ACCATTTCGA

651  GTTCGAGAT  GACAAGCAAT  TTGCCGAGAT  GCTGCGTCAT  CCGCAGAACT
   CAGCTCTAG  CTGTTGTTA  AACGGTCTA  CGACGCACTA  GGCCTTTGA

701  TCGGCCTGAG  CGACGTCGAG  AACCCCTGCT  ACGACGGCG  CTATGTGTGG
   AGCCGGACTC  GGTGAGCTC  TTGGGACGA  TGCTGCCGC  GATACACACC

751  AAGCCGTTT  CCACCCGAG  CGTCAGCAC  GACCGCCAG  TCTCCGCTT
   TTCCGGCAAC  GGTGGGCTC  GCAGTCGTG  CTGGCGGTG  AGAGGCGGAA

801  CAGTCCGAG  GAACGCTCG  CCATCGCCG  CACCCGCTG  CTGGCAGAG
   CTCAGGCGTC  CTTCCGAGC  GGTAGCGGC  GTTGGGCGAC  GACCGTGTCC

851  CCCTTGCCAG  TCCTATGGC  CGCCGCAAG  CCAGCCCCCT  CAACTCTGAG
   GGCAACGGTC  AGGATACCG  GCGGCTCCG  GGTCCGGGGA  GTTACACTTC

901  GGCAGATGT  TCTGGATCA  GGTACACCC  ACCACTGTG  TGCAGCAGC
   CCGTTCTACA  AGACCTAGT  CCATGTGGC  TGGTGACAG  ACGTGCCTG

951  CCTGAGCGAG  CGCCCGCCA  CCTTCATCA  GACCCAGTAC  GAGTTCCTG
   GCACTCGCTC  GCGCGCGGT  GGAAGTAGCT  CTGGGTCTAG  CTCAGGAGC

1001  CCCACGATG  A
      GGTGCTTAC  T

```

FIGURE 72 (SEQ ID No. 24)

```

1  ATGTTTAAGT TTAABAAAGAA TTTCTTAGTT GGATTATCGG CAGCTTAAAT
   TACAAATTCG AATTTTTCCT AAAGAAATCAA CCTAATAGCC CTCGAAATTA

51  GAGTATTAGC TTGTTTTCGG CAACCGCCTC TGCAGCTAGC GCCGACAGCC
   CTCATAATCG AACAAAAGCC GTTGGCGGAG ACGTCGATCG CGGCTGTGGG

101  GTCCCGCCTT TTCCCGGATC GTGATGTTGG GCGACAGCCT CTCGGATACC
   CAGGGCGGAA AAGGGCCTAG CACATACAAGC CGCTGTGCGA GAGGCTATGG

151  GGCAAAATGT ACAGCAAGAT GCGCGGTTAC CTCCTCTCCA GCCCGCCCTA
   CCGTTTACCA TGTCGTTCCTA CGCGCCAAATG GAGGGGAGGT CCGGCGGGAT

201  CTATGAGGGC CGTTTCTCCA ACGSACCCGT CTGGCTGGAG CAGCTGACCA
   GATACTCCCG GCAAGAGAGT TGCCTGGGCA GACCGACCTC CTCGACTGCT

251  AACAGTTCCC GGGTCTGACC ATCGCCCAAG AAGCGGAAAG CCGTGGCACT
   TTCTCAAGGG CCGAGACTGG TAGCGGTTGC TTGCGCTTCC GCCACGGTGA

301  GCGGTGGGTT ACAACAAGAT CTCTGGAAT CCCAAGTATC AGGTCAATCA
   CGGCACCGAA TGTGTCTCTA GAGGACCTTA GGGTTCATAG TCCAGTAGTT

351  CAACCTGGAC TACGAGGTCA CCCAGTTCTT GCAGAAAGAC AGCTTCAAGC
   GTTGGACCTG ATGCTCCAGT GGGTCAAGAA CGTCTTTCAG TCGAAGTTCC

401  CGGACGATCT GGTGATCCTC TGGGTGGGTG CCAATGACTA TCTGGCCTAT
   GCTGCTAGA CCACTAGGAG ACCCAGCCAC GGTACTGAT AGACCGGATA

451  GGCTGGGACA CAGAGCAGGA TGGCAAGCGG GTTCGGGATG CCATCAGCGA
   CCGACCTTGT GCTCTGCTCT ACGGTTCCGC CAAGCGGTAC GGTAGTGGCT

501  TCGCGCCAAC CGCATGGTAC TGAACGGTGC CAAGCAGATA CTGCTGTTCA
   ACGCCGGTTG GCGTAACATG ACTTGCCACG GTTCGTCTAT GACGACAAAT

551  ACCGTCCCGA TCTGGGCCAG AACCCGTGAG CTGCGAGTCA GAAGCTGCTC
   TGGACGGCCT AGACCCGGTC TTGGGCGATC GAGCCTCAGT CTTCCACCAG

601  GAGGCGGTCA GCCATGCTC CGCCTATCAC AACCGAGTGC TGTGACCTT
   CTCGCCCACT CGGTACAGAG GCGGATAGTG TTGCTCGACC ACGACTTGA

651  GGCACGCCAG CTGGCCCCCA CCGSCATGGT AAAGCTGTTT GAGATCGACA
   CCGTGCAGTC GACCGGGGGT GCGCGTACCA ITTCGACAAG CTCTAGCTGT

701  AGCAATTTCG CGAGATGCTG CGTGATCCGC AAGACTTCGG CTTGAGCGAC
   TCGTTAAAGG GCTCTACGAC GCACTAGGCG TCTTGAAGCC GCACTCGCTG

751  GTGAGAAACC CCGCTACGA CCGCGGCTAT CTGTGGAAGC CATTGCGCAC
   CAGCTCTTGG GGACGATGCT GCGCGCGATA CACACCTTCG GCAAAAGGTC

801  CCGCAGCGTC AGCACCAGCC GCCAGCTCTC CGCCTTCACT CCGCAGGAAC
   GCGGTGCGAG TCGTGGCTGG CCGTCTGAGAG GCGGAAGTCA GCGCTCCTTG

851  GCCTCGCCAT CGCCGGCAAC CCGCTGCTGG CACAGGCGCT TGGCAGTCTT
   CCGAGCGGTA GCGGCGCTTG GCGACGACC GTTTCGCGCA ACGCTCAGGA

901  ATGGCCCGCC GACGCGCCAG CCCCCCAAC TGTGAGGGCA AGATGTTCTG
   TACCGGCGCG CCGCGCGCTC GGGGGAGTTG ACACCTCCGT TCTACAAGAC

951  GGATCAGGTA CACCCGACCA CTGTCGTGCA CGCAGCCCTG AGCGAGCGCG
   CCTATCCAT GTGGGCTGGT GACAGCACGT CGCTCGGAC TCGCTCGCGC

1001 CCGCCACCTT CACGCGGAC CAGTACGAGT TCCTCGCCCA CTGATGA
   GCGGTGGAA GTAGCGCTTG GTCATGCTCA AGGAGCGGGT GACTACT

```