

(19) **DANMARK**

(10) **DK/EP 3071692 T3**



Patent- og
Varemærkestyrelsen

(12) Oversættelse af
europæisk patentskrift

-
- (51) Int.Cl.: **C 12 N 9/42 (2006.01)** **C 12 N 15/56 (2006.01)**
- (45) Oversættelsen bekendtgjort den: **2018-08-13**
- (80) Dato for Den Europæiske Patentmyndigheds bekendtgørelse om meddelelse af patentet: **2018-05-02**
- (86) Europæisk ansøgning nr.: **14821723.5**
- (86) Europæisk indleveringsdag: **2014-11-21**
- (87) Den europæiske ansøgnings publiceringsdag: **2016-09-28**
- (86) International ansøgning nr.: **FR2014052984**
- (87) Internationalt publikationsnr.: **WO2015075391**
- (30) Prioritet: **2013-11-22 FR 1361509**
- (84) Designerede stater: **AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR**
- (73) Patenthaver: **IFP Energies nouvelles, 1 & 4 avenue de Bois-Préau, 92500 Reuil-Malmaison, Frankrig**
Proteus, ZI de la Vigne aux Loups, 23 rue Bossuet, 91160 Longjumeau, Frankrig
Centre National de la Recherche Scientifique (CNRS), 3, rue Michel Ange, 75794 Paris Cedex 16, Frankrig
- (72) Opfinder: **MARGEOT, Antoine, 212 rue Marcadet, F-75018 Paris, Frankrig**
BENOIT, Yves, 10 rue Carnot, F-95480 Pierrelaye, Frankrig
PERSILLON, Cécile, 3 rue de la comtesse, F-30000 Nîmes, Frankrig
AYRINHAC, Céline, 7 lotissement le Coteau, F-30350 Domessargues, Frankrig
ULLMANN, Christophe, 5 rue Deparcieux, F-30000 Nîmes, Frankrig
BONZOM, Olivier, 5 Rue de la Servie, F-30000 Nîmes, Frankrig
FORT, Sébastien, 55 Chemin du Parc, F-38410 Vaulnaveys-le-Haut, Frankrig
PETIT, Maud, 14 Rue Jean Baptiste Ducrocq, F-59700 Marcq en Baroeul, Frankrig
LENON, Marine, 27 Chemin de Vinay, F-38360 Sassenage, Frankrig
ARMAND, Sylvie, 14 rue Charrel, 38000 Grenoble, Frankrig
- (74) Fuldmægtig i Danmark: **Plougmann Vingtoft A/S, Rued Langgaards Vej 8, 2300 København S, Danmark**
- (54) Benævnelse: **ENDOGLUCANASE-VARIANTER MED FORBEDRET AKTIVTET, OG ANVENDELSER AF SAMME**
- (56) Fremdragne publikationer:
WO-A1-2008/148131
WO-A1-2010/088387
WO-A1-2011/057140
WO-A1-2011/059740
WO-A1-2013/028701
WO-A2-2009/085935
WO-A2-2011/153516
WO-A2-2012/036810

Fortsættes ...

WO-A2-2012/044915

GUNSELI BAYRAM AKCAPINAR ET AL: "From in silico to in vitro: Modelling and production of *Trichoderma reesei* endoglucanase 1 and its mutant in *Pichia pastoris*", JOURNAL OF BIOTECHNOLOGY, vol. 159, no. 1-2, 1 mai 2012 (2012-05-01) , pages 61-68, XP055039736, ISSN: 0168-1656, DOI: 10.1016/j.jbiotec.2012.01.001

Description

The possibility of producing ethanol from cellulose has received much attention because of the availability of large quantities of raw material and the interest of ethanol as a fuel. The cellulosic natural raw materials for such a process are designated by the term "biomass". Numerous types of biomass, for example wood, agricultural residues, herbaceous cultures and municipal solid waste, have been considered as potential raw materials for the production of biofuel. These materials mainly consist of cellulose, hemicellulose and lignin.

Cellulose is a polymer consisting of molecules of glucose connected by beta 1-4 bonds, which are very resistant to degradation and depolymerisation. Once the cellulose has been converted into glucose, the latter is easily fermented into biofuel, for example ethanol, by using a yeast.

The oldest methods studied for converting cellulose into glucose are based on acid hydrolysis. This process can take place in the presence of concentrated or diluted acids. However, several disadvantages such as the poor recovery of the acid when concentrated acids are used and the low production of glucose when diluted acids are used have a negative effect on the economy of the acid hydrolysis process.

To overcome the disadvantages of the acid hydrolysis process, the processes of conversion of cellulose have more recently focused on enzymatic hydrolysis, using enzymes such as cellulase. This enzymatic hydrolysis of the lignocellulosic biomass (for example cellulose) has, however, the disadvantage of being a costly industrial process. It is therefore necessary to use strains of cellulase-secreting microorganisms that have ever-better performance. In this respect, many microorganisms comprise enzymes that hydrolyse cellulose, such as the fungi *Trichoderma*, *Aspergillus*, *Humicola*, *Fusarium* and bacteria such as *Thermomonospora*, *Bacillus*, *Cellulomonas* and *Streptomyces*. The enzymes secreted by these microorganisms have three types of activities useful in the conversion of the cellulose into glucose and are divided into three groups: endoglucanases, which attack the fibres of celluloses randomly and internally, the exoglucanases that attack the ends of the fibres while freeing cellobiose, and the β -glucosidases that hydrolyse this cellobiose into glucose. Other classes of enzymes such as hemicellulases or the recently discovered class of enzymes of the polysaccharide mono-oxygenases can also play a role in the efficiency of the hydrolysis.

There is a strong industrial interest in reducing the cost of enzymatic hydrolysis, and this reduction involves the use of a reduced dose of enzymes and thus of more efficient enzyme cocktails. Consequently, a plurality of patent applications describe natural enzymes with greater capacities than those of *Trichoderma reesei*, or variants improved via genetic engineering. Examples include the patent applications US2010304464, WO2010066411 and WO2013029176 relating to the exoglucanases, the applications WO2007109441, WO2012149192, WO2012036810 and WO2010076388 relating to the

endoglucanases, the applications WO2010029259, WO2010135836, WO2012044915, WO2011057140 and WO2010022518 relating to the beta-glucosidases, or the applications WO12135659, WO12149344 relating to the polysaccharide mono-oxygenases.

The application WO2011153516 also describes strains of yeasts that express saccharolytic enzymes.

The application WO2013028701 also describes variants of glycoside hydrolase GHG61.

The application WO2008141131 describes proteins having an improved cellulolytic activity.

Gunseli Bayram Akcapinar et al. describe the production of endonuclease I from *Trichoderma reesei* in the Journal of Biotechnology, vol. 159, n°1-2, 1 May 2012, pages 61-68.

The enzymes that hydrolyse lignocellulosic biomass are classified in the CAZy system (Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V., & Henrissat, B. (2009). The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. Nucleic acids research, 37, D233-8) on the basis of mainly structural criteria. The endoglucanases can belong to the families GH 5, 6, 7, 8, 9, 12, 16, 18, 19, 26, 44, 45, 48, 51, 74 and 124.

In order for hydrolysis of the lignocellulosic biomass to be efficient and economically profitable, the enzymatic mixture must comprise balanced proportions of enzymes having various enzymatic activities, *inter alia*, but not exclusively, such as exoglucanases, endoglucanases, xylanases and β -glucosidases. For example, in the native mixtures of *Trichoderma reesei*, the presence of 60-70% exoglucanases, 15-20% endoglucanases, a few percent of hemicellulase and approximately 5-10% β -glucosidases is generally observed. This mixture is suitable for hydrolysing the majority of pretreated substrates (typical ex. wheat straw steam-exploded under acidic conditions) with acceptable yields. In summary, the increase in the endoglucanase activity must not occur at the expense of the other enzymatic activities. The functional specificities of these enzymes are not well known at present. The genome of *Trichoderma reesei* comprises at least 3 main enzymes, from the families 7 (EG1, cel7b), 5 (EG2, cel5a) and 12 (EG3, Cel12a). The enzymes EG1 and EG2 are the majority endoglucanases and can represent up to 10-20% by mass of the complete enzyme cocktail produced by *T. reesei*.

The endoglucanases (EC 3.2.1.4), the first enzymes to act on the cellulose, are known to have a major role in the hydrolysis by increasing the number of sites that the exoglucanases can attack while reducing the degree of polymerisation of the attacked microfibrils. Recent works (Szijártó, N., Siika-aho, M., Sontag-Strohm, T., & Viikari, L. (2011). Liquefaction of hydrothermally pretreated wheat straw at high-solids content by purified *Trichoderma* enzymes. *Bioresource technology*, 102(2), 1968-74) highlight their role in the reduction of the viscosity of the biomass during the first hours of the hydrolysis. This reduction in viscosity can have a very significant impact on the operating costs of the method.

The problem of viscosity is exacerbated in the case of methods that require the use of a low temperature such that the saccharification simultaneous to the fermentation (SSF), which involves both

the enzymes that hydrolyse the biomass and the microorganism that converts the monomers of sugars into ethanol.

The hydrolysis and the fermentation can be carried out using various blueprints. The most common involves separate hydrolysis and fermentation (SHF – Separate Hydrolysis and Fermentation).

5 This method allows each step to be optimised by maintaining optimal reaction conditions. This fermentation is carried out extemporaneously, at a temperature between approximately 28°C and approximately 30°C, while the hydrolysis generally takes place at a temperature of at least 45°C. However, during SHF, the sugars freed at the end of the reaction are present at a very high concentration and lead to an inhibition of the enzymes, slowing down the efficiency of the method. To overcome these
10 disadvantages, another type of method is possible. In SSF, the two steps (hydrolysis and fermentation of the hexoses) take place simultaneously, preventing the accumulation of the sugars at concentrations that inhibit the enzymes. The investment costs are also reduced due to the use of a single reactor. The hydrolysis rate is greater due to the lack of inhibition since the sugars freed are immediately used for the fermentation into ethanol. In this method, the temperature of the reactor necessarily represents a
15 compromise between the optimal temperatures of hydrolysis and of fermentation, typically between approximately 30°C and approximately 35°C. However, at such a temperature, the activity of the cellulolytic enzymes is reduced by approximately 30%.

SSF also allows the expression of enzymes that degrade the cellulose in the organism that ferments the sugars, which limits, or in an extreme case eliminates the need to use enzymes produced
20 during a separate step. However, producing large quantities of enzymes with the fermenting organisms and thus obtaining a significant activity can be problematic and limits the viability of these approaches.

Consequently, obtaining enzymes that maintain an efficient endoglucanase activity at the optimal temperatures of hydrolysis and of fermentation (or between 30°C and 50°C) while keeping the proportions of all of the enzymes of the mixture would represent a significant improvement for the
25 method for converting lignocellulosic biomass into biofuel.

Description

The inventors have developed a polypeptide having an improved endoglucanase activity, in particular with respect to the endoglucanase activity of the wild-type protein EG1 having sequence SEQ
30 ID NO:2. EG1 corresponds to the endoglucanase 1 of *Trichoderma reesei*.

In this respect, the applicants, should be given a lot of credit for finding, after much research, an isolated or purified polypeptide having an improved endoglucanase activity with respect to the endoglucanase activity of the reference protein EG1 (SEQ ID NO:2).

The description thus relates to a polypeptide chosen from the group consisting of:

i) a sequence of amino acids chosen from SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12; SEQ ID NO:14 SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:20, SEQ ID NO:22, SEQ ID NO:24 and SEQ ID NO:26;

5 ii) a sequence of amino acids having a percent identity of at least 70%, preferably 75%, 80%, 85%, 90%, 95%, 98% or 99%, with respect to the sequence SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12; SEQ ID NO:14 SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:20, SEQ ID NO:22, SEQ ID NO:24 or SEQ ID NO:26.

The invention is defined in the claims.

10 Preferably, the polypeptide according to the invention is characterised in that its expression in a fermenting organism is at least equal to the expression of the reference protein EG1 (SEQ ID NO:2).

The percent identity of a given sequence with respect to SEQ ID NO:4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24 or 26, corresponds to the number of identical residues between this given sequence and SEQ ID NO:4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24 or 26 divided by the number of residues in SEQ ID NO:4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24 or 26. When the GenomeQuest database is used, said percent identity with
15 respect to SEQ ID NO: 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24 or 26 corresponds to the Query percent identity (% id Query), where Query corresponds to the sequence SEQ ID NO:4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24 or 26.

A person skilled in the art can for example determine the increase or in other words the improvement in the enzymatic activity either using the substrate Carboxy-Methyl-Cellulose (CMC) or
20 with a chromogenic substrate (p-Nitrophenyl glycoside). The enzymatic activity is revealed, respectively, via colorimetric assay of the reducing sugars or of the freed nitrophenol.

Preferably, the polypeptide of the invention has an enzymatic activity improved by at least 10%, preferably by at least 20%, preferably by at least 30%, with respect to the endoglucanase activity of the protein EG1 having the sequence of amino acids SEQ ID NO:2.

25 An example of a protocol, which a person skilled in the art can use to determine whether a polypeptide according to the invention has an improved enzymatic activity with respect to that of the reference protein EG1 (SEQ ID NO:2), is the following:

- formation of a stock culture of *E. coli* expressing a polypeptide according to the invention overnight at 37°C;

30 - inoculation of an LB culture medium with 1% stock culture at 37°C until an optical density of 0.4 is obtained;

- culture of said cells at 20°C for 18h;

- centrifugation for 5 minutes at 7900rpm;

- resuspension of the cell pellets with 100mM citrate phosphate buffer at pH 5 containing
35 1mg/mL of lysozyme (final OD₆₀₀ of 100);

- incubation of the resuspended cells for 30 minutes over ice;
- lysis of the cells via 3 cycles of freezing/thawing;
- fractionation of the DNA via sonication;
- centrifugation for 30 minutes at 13000rpm;
- incubation of 100μL of breaking supernatant with 100μL of 100mM citrate phosphate buffer at pH 5 containing 1% CMC for 6h at 35 and 50°C;
- removal of 100μL of reaction;
- addition of 100μL of DNS reactant (Miller, 1959);
- incubation for 5 minutes at 100°C;
- incubation for 3 minutes over ice;
- centrifugation for 10 minutes at 3000rpm;
- reading of the optical density at 540nm on 150μL of supernatant.

The object of the invention is also a purified or isolated nucleic acid coding at least one polypeptide according to the invention as described above. TABLE 1 below comprises the identifications of the nucleic and peptide sequences for EG1 of *T. reesei* ("wild-type"), the putative endoglucanases of *Chaetomium globosum* (C) and of *Aspergillus fumigatus* (A), as well as for the polypeptides and nucleotides of the invention and of the description.

According to the description, said purified or isolated nucleic acid can be chosen from the following sequences: SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11; SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:23 and SEQ ID NO:25.

TABLE 1

Clones	Nucleic acid	Polypeptide
EG1 (wild-type)	SEQ ID NO:1	SEQ ID NO:2
76B4	SEQ ID NO:3	SEQ ID NO:4
105F11	SEQ ID NO:5	SEQ ID NO:6
107H12	SEQ ID NO:7	SEQ ID NO:8
154E4	SEQ ID NO:9	SEQ ID NO:10
202C12	SEQ ID NO:11	SEQ ID NO:12
272A9	SEQ ID NO:13	SEQ ID NO:14
278F10	SEQ ID NO:15	SEQ ID NO:16
293B2	SEQ ID NO:17	SEQ ID NO:18
309A11	SEQ ID NO:19	SEQ ID NO:20
11G8	SEQ ID NO:21	SEQ ID NO:22
92A12	SEQ ID NO:23	SEQ ID NO:24
240H12	SEQ ID NO:25	SEQ ID NO:26

Gene C	SEQ ID NO:27	SEQ ID NO:28
Gene A	SEQ ID NO:29	SEQ ID NO:30

The invention also relates to a vector comprising a nucleic acid according to the invention as described above.

According to the invention, “vector” means any sequence of DNA into which it is possible to insert fragments of foreign nucleic acid, the vectors allowing foreign DNA to be introduced into a host cell. A non-exhaustive list of examples of vectors includes: plasmids, cosmids, yeast artificial chromosomes (YAC), bacterial artificial chromosomes (BAC), P1-derived artificial chromosomes (PAC) and vectors derived from viruses.

According to the invention, the nucleic acid as described above can be operationally linked to a promoter, a terminator or any other sequence necessary for its expression in the host cell.

The vector according to the invention can also carry a selection marker. A “selection marker” means a gene, the expression of which gives the cells that contain it a characteristic allowing them to be selected. This involves for example a gene for resistance to antibiotics.

The object of the invention is also an isolated host cell comprising at least one of the polypeptides according to the invention as described above, at least one of the nucleic acids according to the invention as described above, or at least one of the vectors according to the invention as described above.

A person skilled in the art can introduce one of the polypeptides, one of the nucleic acids or one of the vectors according to the invention as described above into the host cell according to the invention via well-known conventional methods. For example, these include treatment with calcium chloride, electroporation, the use of a particle gun.

According to one embodiment, a person skilled in the art can introduce, into the host cell according to the invention and via conventional methods, a plurality of copies of a nucleic acid according to the invention coding a polypeptide according to the invention having an improved endoglucanase activity according to the invention.

According to one embodiment, the isolated host cell according to the invention as described above is chosen from *Trichoderma*, *Aspergillus*, *Neurospora*, *Humicola*, *Myceliophthora*, *Chrysosporium*, *Penicillium*, *Fusarium*, *Thermomonospora*, *Bacillus*, *Pseudomonas*, *Escherichia*, *Clostridium*, *Cellulomonas*, *Streptomyces*, *Yarrowia*, *Pichia* and *Saccharomyces*.

According to a preferred embodiment, the isolated host cell according to the invention as described above is chosen from *Trichoderma reesei*, *Trichoderma viridae*, *Trichoderma koningii*, *Aspergillus niger*, *Aspergillus nidulans*, *Aspergillus wentii*, *Aspergillus oryzae*, *Aspergillus phoenicis*, *Myceliophthora thermopila*, *Chrysosporium lucknowense*, *Neurospora crassa*, *Humicola grisea*,

Penicillium pinophilum, *Penicillium oxalicum*, *Escherichia coli*, *Clostridium acetobutylicum*, *Clostridium saccharolyticum*, *Clostridium benjerinckii*, *Clostridium butylicum*, *Pichia pastoris*, *Yarrowia lipolytica* and *Saccharomyces cerevisiae*.

5 According to a preferred embodiment, the isolated host cell according to the invention as described above is chosen from *Trichoderma reesei* and *Saccharomyces cerevisiae*.

The object of the invention is also the use of any one of the polypeptides according to the invention described above for the hydrolysis of cellulose.

The object of the invention is also the use of any one of the polypeptides according to the invention described above for the production of biofuel.

10 According to the invention, the term biofuel can be defined as any product that results from the transformation of biomass and can be used for energy purposes. Without imposing any limitations, examples of this include biogases, products that can be incorporated (optionally after a later transformation) into a fuel or can be a fuel on their own, such as alcohols (ethanol, butanol and/or isopropanol according to the type of fermenting organism used), solvents (acetone), acids (butyric),
15 lipids and their derivatives (short-chain or long-chain fatty acids, fatty acid esters), and hydrogen.

Preferably, the biofuel according to the invention is an alcohol, for example ethanol, butanol and/or isopropanol. More preferably, the biofuel according to the invention is ethanol.

In another embodiment, the biofuel is biogas.

In another embodiment, the product is a molecule of interest to the chemical industry, for
20 example such as, another alcohol such as 1,2-propane diol, 1,3-propane diol, 1,4-butane diol, 2,3-butane diol, organic acids such as acetic, propionic, acrylic, butyric, succinic, malic, fumaric, citric, itaconic acid, or hydroxyacids such as glycolic, hydroxypropionic or lactic acid.

An embodiment for producing an enzyme cocktail useful for the hydrolysis of lignocellulose is described below.

25 The strains of filamentous fungi, preferably *Trichoderma*, more preferably *T. reesei*, capable of expressing at least one polypeptide according to the invention are cultivated in fermenters, in the presence of a carbon substrate, such as lactose or glucose, chosen for the growth of the microorganism. In one embodiment, this carbon substrate, according to its nature, is introduced into the fermenter before sterilisation or is sterilised separately and introduced into the fermenter after sterilisation of the
30 latter in order to obtain an initial concentration of 20 to 35g/L.

An aqueous solution containing the substrate chosen for the production of the enzymes is then added. An enzymatic composition acting on the lignocellulosic biomass produced by the fungi is finally recovered via filtration of the culture medium. This composition contains in particular the β -glucosidase, the exoglucanase and the endoglucanase according to the invention.

In one embodiment, the aqueous solution containing the substrate chosen for the production of the enzymes is prepared at a concentration of 200-250g/L. This solution further contains preferably an inductive substrate such as lactose. This aqueous solution is injected after the initial carbon substrate is used up in such a way as to provide an optimised quantity, between 35 and 45mg/g of cells ("fed batch"). During this "fed batch" phase, the residual concentration of sugar in the culture medium is less than 1g/L and the enzymes acting on the lignocellulosic biomass are secreted by the fungus. Said enzymes can be recovered via filtration of the culture medium.

The object of the invention is an enzymatic composition suitable for acting on the lignocellulosic biomass, said enzymatic composition being produced by filamentous fungi and comprising at least one polypeptide having an improved endoglucanase activity with respect to the endoglucanase activity of the reference protein EG1. "Filamentous fungi" means in particular *Trichoderma*, more preferably *T. reesei*.

Finally, the object of the invention is a method for producing biofuel from the biomass, comprising the following successive steps:

- placing the biomass to be hydrolysed in suspension in aqueous phase;
- hydrolysis, in the presence of an enzymatic composition, of the lignocellulosic biomass as described above in such a way as to produce a hydrolysate containing glucose;
- fermentation of the glucose of the hydrolysate in such a way as to produce a fermentation must;
- separation of the biofuel from the fermentation must.

In one embodiment, the biomass to be hydrolysed is placed in suspension in aqueous phase at 6 to 40% dry matter, preferably 20 to 30%. The pH is adjusted to between 4 and 5.5; preferably between 4.8 and 5.2 and the temperature between 40 and 60°C, preferably between 45 and 50°C. The hydrolysis reaction is started by the addition of the enzymatic composition acting on the lignocellulosic biomass; the quantity usually used is from 10 to 30 mg of excreted proteins per gram of pretreated substrate or less. The reaction generally lasts from 15 to 48 hours. The reaction is followed by assay of the freed sugars, in particular the glucose. The solution of sugars is separated from the non-hydrolysed solid fraction, substantially consisting of lignin, via filtration or centrifugation and then treated in a fermentation unit.

After the step of fermentation, the biofuel is separated from the fermentation must for example via distillation.

Another object of the invention is a method for producing biofuel from biomass, characterised in that it comprises the following successive steps:

- placing the biomass to be hydrolysed in suspension in aqueous phase;

- simultaneous addition of an enzymatic composition as defined above and of a fermenting organism in such a way as to produce a fermentation must;
- separation of the biofuel from the fermentation must.

Preferably, the enzymatic composition and the fermenting organism are added simultaneously and then incubated at a temperature between 30°C and 35°C in order to produce a fermentation must.

According to this embodiment, the cellulose present in the biomass is converted into glucose, and at the same time, in the same reactor, the fermenting organism (for example a yeast) converts the glucose into final product according to an SSF method (Simultaneous Saccharification and Fermentation) known to a person skilled in the art. According to the metabolic and hydrolytic capabilities of the fermenting organism, correctly carrying out the operation may require the addition of a more or less significant quantity of exogenous cellulolytic mixture.

In another embodiment, the fermenting organism produces the polypeptide that is the object of the invention via secretion or on the surface of its cell, optionally together with other enzymes acting on the lignocellulosic biomass, thus limiting or eliminating the need for enzymes produced by the filamentous fungus. Preferably, the fermenting organism is a host cell as described above.

Thus, preferably, the object of the invention is a method for producing biofuel from the biomass, comprising the following successive steps:

- placing the biomass to be hydrolysed in suspension in aqueous phase;
- adding one or more host cells as described above, with a fermenting organism and/or an enzymatic composition as described above, in such a way as to produce a fermentation must;
- separating the biofuel from the fermentation must.

Preferably, the host cells with the enzymatic composition and/or the fermenting organism, are added and then incubated at a temperature between 30°C and 35°C in order to produce a fermentation must.

The use of the polypeptide having a better endoglucanase activity according to the present invention thus has the advantage of obtaining a better yield in terms of glucose production while using less enzyme than before, which also has an economic advantage.

Other aspects, objects, advantages and features of the invention, will be presented upon reading the non-limiting description that follows and that describes preferred embodiments of the invention given via examples and the DRAWINGS.

FIGURE 1 is a graph showing the hydrolysis of the 1% CMC by the reference endoglucanase (EG1) and its mutant (154E4) secreted into the culture medium by the strains Sca-EG1 and Sca-154E4, respectively.

FIGURE 2 is a graph presenting the results of SHF for the cocktail resulting from the strain 154E4/8 (SEQ ID NO:10), a reference cocktail produced by the strain CL847 ΔEG1 (ΔEG1) and

supplemented with β -glucosidase, and another reference cocktail produced by the strain CL847 Δ EG1 retransformed with the reference gene EG1 (Δ EG1cEG1), supplemented with β -glucosidase.

FIGURE 3 is a graph presenting the results of SHF for the cocktail resulting from the strain 11G8/10 (SEQ ID NO:22), a reference cocktail produced by the strain CL847 Δ EG1 (Δ EG1) and supplemented with β -glucosidase, and another reference cocktail produced by the strain CL847 Δ EG1 retransformed with the reference gene EG1 (Δ EG1cEG1), supplemented with β -glucosidase.

FIGURE 4 is a graph presenting the results of SSF for the two cocktails resulting from the strains 154E4/2 154 E4/8 (SEQ ID NO:10), a reference cocktail produced by the strain CL847 Δ EG1 (Δ EG1) and supplemented with β -glucosidase, and another reference cocktail produced by the strain CL847 Δ EG1 retransformed with the reference gene EG1 (Δ EG1cEG1), supplemented with β -glucosidase.

FIGURE 5 is a graph presenting the results of SSF for the three cocktails resulting from the strains 11G8/10, 11G8/12 and 11G8/13 (SEQ ID NO:22), a reference cocktail produced by the strain CL847 Δ EG1 (Δ EG1) and supplemented with β -glucosidase, and another reference cocktail produced by the strain CL847 Δ EG1 retransformed with the reference gene EG1 (Δ EG1cEG1), supplemented with β -glucosidase.

EXAMPLES

EXAMPLE 1: 1st round of L-shuffling

The sequence of the reference gene EG1 from *Trichoderma reesei* (SEQ ID NO:1) was subjected to a first round of L-shuffling according to the method described in EP1104457B1 with the genes of the putative endoglucanase of *Chaetomium globosum* (SEQ ID NO:29) and of the endoglucanase of *Aspergillus fumigatus* (SEQ ID NO:27) each having approximately 60% identity with the reference gene SEQ ID NO:1.

1-High-throughput screening

A high-throughput screening test was developed in order to select the best clones resulting from the L-shuffling, that is to say, those having at least 20% improvement in the endoglucanase activity with respect to the reference enzyme SEQ ID NO:2.

The high-throughput screening test was carried out according to the following steps:

- isolation, on agar, of the clones of *E. coli* expressing the L-shuffling variants of the recombinant enzyme according to the invention and placement of said colonies into precultures in an LB medium overnight at 37°C;

- inoculation of an LB medium at 6% with the preculture, then incubation for 5 hours at 37°C, then 17 hours at 20°C;

- centrifugation for 10 minutes at 3000rpm;

- lysis of the cells via addition of 80µL of a solution of lysozyme at 1mg/mL in 0.1M citrate phosphate buffer at pH 5;

- incubation for 4 hours at ambient temperature;

- addition of 80µL of pH 5 0.1M citrate phosphate buffer containing 1% carboxy-methyl cellulose;

- incubation for 17 hours at 35°C;

- centrifugation for 10 minutes at 3000rpm;

- removal of 100µL of supernatant;

- addition of 100µL of DNS reagent;

- incubation for 10 minutes at 100°C and then 5 minutes over ice;

- reading of the OD at 540nm on 120µL.

Under these high-throughput screening conditions, an improvement in the endoglucanase activity (increase in the OD at 540nm) with respect to the reference enzyme EG1 (SEQ ID NO:2) was found in a plurality of clones, including in particular the clones 76B4,105F11, 107H12, 154E4, 202C12, 272A9, 278F10, 293B2 and 309A11.

2-Determining the improvement in the endoglucanase activity

2-1/ On the substrate carboxy-methyl cellulose (CMC)

In order to estimate the kcat of the variants selected in the first round of L-shuffling with respect to the reference enzyme (SEQ ID NO:2), the following is carried out:

- preparation of a stock culture of *E. coli* expressing a recombinant enzyme according to the invention overnight at 37°C;

- inoculation of an LB culture medium with 1% stock culture at 37°C until an optical density of 0.4 at 600nm is obtained;

- culture of said cells at 20°C for 18 hours;

- centrifugation for five minutes at 7900rpm;

- resuspension of the cell pellets with 0.1M citrate phosphate buffer at pH 5 containing 1mg/mL of lysozyme (final OD₆₀₀ of 100);

- incubation of the resuspended cells for 30 minutes over ice;

- lysis of the cells via 3 cycles of freezing/thawing;

- fractionation of the DNA via sonication for 3 seconds at power 5;

- centrifugation for 30 minutes at 13000rpm;

- incubation of 100µL of breaking supernatant with 100µL of 0.1M citrate phosphate buffer at pH 5 containing 1% CMC for 6 hours at 35 and 50°C;

- removal of 100 μ L of reaction;
- addition of 100 μ L of DNS reagent;
- incubation for 5 minutes at 100°C;
- incubation for 3 minutes over ice;
- centrifugation for 10 minutes at 3000rpm;
- reading of the optical density at 540nm on 150 μ L.

According to the invention, the calculation of the kcats is carried out in the following manner:

- expression of the ODs at 540nm according to the quantity of protein of interest (in nM);
- subtraction of the value of the negative control;
- division by the coefficient of the reference range for glucose (various quantities of glucose are revealed with the DNS);
- division by the reaction time (360 minutes).

TABLE 2 presents the value for the kcats and the improvement factor obtained for the clones 76B4, 105F11, 107H12, 154E4, 202C12, 236C11, 272A9, 278F10, 293B2 and 309A11 with respect to the reference protein EG1 (SEQ ID NO:2) under the experimental conditions of the test of activity on CMC.

TABLE 2: Endoglucanase activity on CMC

		35°C		50°C	
	Clone	Kcat (min ⁻¹)	Improvement factor	Kcat (min ⁻¹)	Improvement factor
Clones of the first round	76B4	0.25	1	0.35	1.1
	105F11	1.13	4.7	1.24	4
	107H12	1.06	4.4	1.14	3.7
	154E4	1.27	5.3	3.39	10.9
	202C12	0.73	3	0.61	2
	236C11	0.12	5.3	0.05	0.2
	272A9	0.99	4.1	2.33	7.5
	278F10	0.95	4	1.06	3.4
	293B2	1.04	4.3	1.78	5.7
	309A11	1.07	4.5	0.6	1.9
Reference protein	EG1	0.24	1	0.31	1

The results show, for these clones, an improvement in enzymatic activity with respect to the reference enzyme SEQ ID NO:2.

2-2/ On the phosphoric acid swollen cellulose (PASC) substrate

The improvement in activity for the clones 76B4, 105F11, 107H12, 154E4, 202C12, 236C11, 272A9, 278F10, 293B2 and 309A11 was then confirmed on a second substrate: phosphoric acid swollen cellulose (PASC).

The determination of the kcat on this substrate is carried out according to the same protocol as described above. The CMC substrate is replaced by the PASC substrate at the same concentration.

Table 3 presents the value for the kcats and the improvement factor obtained for the clones 76B4, 105F11, 107H12, 154E4, 202C12, 236C11, 272A9, 278F10, 293B2 and 309A11 with respect to the reference protein EG1 (SEQ ID NO:2) under the experimental conditions of the test of activity on PASC.

TABLE 3: Endoglucanase activity on PASC

		35°C		50°C	
	Clone	Kcat (min ⁻¹)	Improvement factor	Kcat (min ⁻¹)	Improvement factor
Clones of the first round	76B4	0.0116	1.66	0.0127	1.81
	105F11	0.0099	1.41	0.0109	1.56
	107H12	0.0069	0.99	0.0068	0.97
	154E4	0.0101	1.44	0.0104	1.49
	202C12	0.0098	1.4	0.0097	1.39
	236C11	0.0102	1.46	0.0103	1.47
	272A9	0.0103	1.47	0.0099	1.41
	278F10	0.0094	1.34	0.0091	1.3
	293B2	0.0089	1.27	0.0088	1.26
	309A11	0.0102	1.46	0.0089	1.27
Reference protein	EG1	0.007	1	0.007	1

These results show an improvement in the enzymatic activity with respect to the reference enzyme EG1 (SEQ ID NO:2).

Example 2: 2nd round of L-shuffling

The improved genes 105F11, 154E4, 202C12, 272A9, 278F10 and 309A10 (respectively SEQ ID NO:5, 9, 11, 13, 15, 19) obtained in the first round of evolution were then subjected to a second round of L-shuffling (again according to the patented method described in EP1104457B1). In order to promote reconstruction on a backbone of the sequence of *Trichoderma*, the reference gene EG1 (SEQ ID NO:1) was reintroduced as a parent gene for the second round of L-shuffling.

1-High-throughput screening

A high-throughput screening test as described above was carried out on the clones obtained after this second round of L-shuffling in order to select the best ones. The activity test was reduced to 2 hours (compared to 17 hours for the screening of the clones resulting from the first round of evolution) in order to take into account the improvements obtained with the first round of L-shuffling.

The activity of the clones generated is compared to the activity obtained with the clone 154E4. This clone, resulting from the first round of L-shuffling, is that which provided the best improvement in activity.

Under these screening conditions, an improvement in the endoglucanase activity with respect to the reference clone 154E4 (SEQ ID NO:10) was found in a plurality of clones, including in particular the clones 11G8 and 240H12.

2- Determination of the improvement in the endoglucanase activity

2-1/ On the substrate carboxy-methyl cellulose (CMC)

In order to determine the kcat, the activities of the clones 11G8, 92A12 and 240H12 were measured via the activity test as described above. The duration of the activity test was reduced to one hour of incubation with the substrate in order to take into account the improvements of these clones.

TABLE 4 presents the values for the kcat and the improvement factor obtained for the clones 11G8, 92A12 and 240H12 with respect to the clone 154E4 (SEQ ID NO:10) under these experimental conditions.

TABLE 4: Endoglucanase activity on CMC

		35°C		50°C	
	Clone	Kcat (min ⁻¹)	Improvement factor	Kcat (min ⁻¹)	Improvement factor
	11G8	804.7	1.8	944.2	1.9
	92A12	420	0.94	478.5	0.95

Clones of the second round	240H12	488.6	1.1	590.8	1.2
Reference protein	154E4	446.4	1	501.3	1

The results show an improvement in activity with respect to the reference clone 154E4 for the clones 11G8 and 240H12.

5 2-2/ On the phosphoric acid swollen cellulose (PASC) substrate

The improvement in activity for the clones 11G8, 92A12 and 240H12 was then confirmed on a second substrate: phosphoric acid swollen cellulose (PASC).

In order to determine the kcat, the activity of these clones was measured at 35 and 50° C via the activity test as described above with PASC as the substrate.

10 TABLE 5 presents the values for the kcat and the improvement factor obtained for the clones 11G8, 92A12 and 240H12 with respect to the clone 154E4 (SEQ ID NO:10) under these experimental conditions.

TABLE 5: Endoglucanase activity on PASC

		35°C		50°C	
	Clone	Kcat (min ⁻¹)	Improvement factor	Kcat (min ⁻¹)	Improvement factor
Clones of the second round	11G8	1.37	0.8	1.72	1.1
	92A12	1.36	0.81	1.7	1.04
	240H12	1.87	1.1	1.9	1.2
Reference protein	154E4	1.68	1	1.63	1

15

The activity of the clone 11G8 is not improved at 35°C on this substrate with respect to the reference clone 154E4. The activity of the clone 11G8 is improved, however, at 50°C. The activity of the clone 240H12 is improved at 35°C and at 50° C.

20 EXAMPLE 3: Cloning of a variant endoglucanase 1 resulting from the first round of L-shuffling in a strain *T. reesei* CL847 ΔEG1

The construction of the DNA fragment to be inserted into *T. reesei*, containing the clone 154E4 resulting from the first round of the L-shuffling, was carried out via PCR fusion. The fragment having a length of approximately 5.4kb consisted of the phleomycin-resistance gene, the sequence coding for the clone 154E4 under control of the promoter and followed by the *cbh1* terminator. Likewise, the reference gene EG1 (SEQ ID NO:1) of *T. reesei* was amplified and fused between the *cbh1* promoter and terminator, resulting in a second construct.

Protoplasts of a strain *T. reesei* CL847 Δ EG1 were transformed according to a conventional method known to a person skilled in the art, via calcium shock and PEG, with 5 μ g of each construct, namely the fragment of DNA containing either the gene 154E4 or EG1. The transformants were selected over a PDA/sucrose selective medium containing 30 μ g/L of phleomycin. Fourteen clones of each transformation were subcultured. After three subculturings in order to isolate pure clones, seven clones having integrated the native gene and five clones having integrated the variant 154E4 and secreting a level of proteins comparable to the strain CL847, were finally obtained.

1-Screening using a test of activity on carboxymethyl cellulose (CMC)

The 11 clones were cultured in 24-well plates containing the following medium:

800 μ L 85% H₃PO₄, 4.2g (NH₄)₂SO₄, 0.3g MgSO₄·7H₂O, 1.5g CornSteep, 1mL Oligo Ferment, 11.6g maleic acid, 10g Solka-Floc and 20g lactose per litre of medium. The pH is adjusted to 5.8-6.0. After 5 days of culture at 30°C, the supernatant is removed and the equivalent of 10mg/L of proteins (measured by the Lowry method) is used for a test of activity on CMC. 150 μ L of a 2% CMC solution in pH 4.8 50mM citrate buffer are mixed with 150 μ L of citrate buffer containing 10mg/L of proteins. The reaction is incubated at 50°C or 35°C for 10 minutes and then inactivated in a boiling water bath. After centrifugation for 5 minutes, 20 μ L of supernatant are removed in order to carry out an assay of reducing sugars using 3,5-dinitrosalicylic acid (DNS). The reduction of the DNS and the formation of 3-amino-5-nitrosalicylic acid are followed by reading of the absorption at 540nm, and the reducing sugars are quantified using a glucose range.

TABLE 6 summarizes the activities obtained with the clones containing the variant 154E4, in comparison to the reference strain CL847, the strain CL847 Δ EG1, and the strain CL847 Δ EG1 retransformed with the reference gene EG1 (Δ EG1cEG1, average obtained with the four best transformants).

TABLE 6: Endoglucanase activity on CMC

Clone	Specific activity	154E4/ Δ EG1cEG1 ratio (50°C)	Specific activity	154E4/ Δ EG1cEG1 ratio (35°C)

	50°C ($\mu\text{mol}/\text{mg}/\text{min}$)		35°C ($\mu\text{mol}/\text{mg}/\text{min}$)	
CL847	12.9 ± 3.1		9.7 ± 0.4	
ΔEG1	6.4 ± 0.4		—	
$\Delta\text{EG1cEG1}$	16.8 ± 2.5		12.1 ± 1.0	
154E4/2	22.5 ± 2.9	1.3	12.7 ± 2.0	1.1
154E4/8	24.1 ± 1.9	1.4	12.5 ± 2.0	1.0
154E4/9	20.2 ± 3.9	1.2	7.4 ± 2.2	0.6

The results show that the activities of the 154E4 variants are between 1.2 and 1.4 times greater than those of the clones retransformed with the reference gene EG1 (SEQ ID NO:1). The improvement is visible at 35°C and at 50°C.

5

2-Cloning of the variant 11G8 obtained after the 2nd round of L-shuffling in a strain *T. reesei* CL847 ΔEG1 and screening of the transformants:

The variant 11G8 was cloned between the promoter and the cbh1 terminator in the plasmid pUT1040 containing a phleomycin-resistance gene as a marker, using BamH1/XhoI double digestion. 5 μg of this vector were used for the transformation of the strain *T. reesei* CL847 ΔEG1 . The transformation of the protoplasts was carried out under the same conditions as for the variant 154E4. At the end of the transformation process and of three successive subculturings carried out with the goal of obtaining pure clones, thirteen clones with a protein production similar to the strain CL847 were obtained and were subjected to screening via measurement of CMCase activity. The activity test is identical to the screening of the clones containing the variant 154E4 from the first round of L-shuffling. Six clones expressing the variant 11G8 show a CMCase activity greater than that of the strain $\Delta\text{EG1cEG1}$. The two best transformants have an activity increased by 70% with respect to the strain expressing the reference gene EG1 (SEQ ID NO:1).

10

15

TABLE 7 summarizes the activities obtained with the clones containing the variant 11G8, in comparison to the reference strain CL847, the strain CL847 ΔEG1 , and the strain CL847 ΔEG1 retransformed with the reference gene EG1 ($\Delta\text{EG1cEG1}$, average obtained with the four best transformants).

20

TABLE 7: Endoglucanase activity on CMC

Clone	Specific activity 50°C ($\mu\text{mol}/\text{mg}/\text{min}$)	11G8/ $\Delta\text{EG1cEG1}$ ratio (50°C)	Specific activity 35°C ($\mu\text{mol}/\text{mg}/\text{min}$)	11G8/ $\Delta\text{EG1cEG1}$ ratio (35°C)
CL847	20.8 ± 0.8		11.8 ± 0.5	
ΔEG1	-		3.3 ± 1.1	
$\Delta\text{EG1cEG1}$	13.2 ± 2.2		9.1 ± 0.7	
11G8/6	15.6 ± 1.8	1.2	10.0 ± 0.4	1.1
11G8/7	17.3 ± 0.7	1.3	10.8 ± 0.9	1.2
11G8/9	17.9 ± 0.4	1.4	11.9 ± 0.2	1.3
11G8/10	15.0 ± 0.3	1.1	13.9 ± 1.6	1.5
11G8/12	17.0 ± 0.1	1.3	15.7 ± 0.8	1.7
11G8/13	17.3 ± 1.2	1.3	15.6 ± 0.4	1.7

The results show that the activities of the 11G8 variants are between 1.1 and 1.7 times greater than those of the clones retransformed with the reference gene SEQ ID NO:1. The improvement is visible at 35°C and at 50°C.

5

Example 4: Recombinant expression of the reference endoglucanase EG1 and of the improved variant 154E4 in *Saccharomyces cerevisiae*

1 - Production of the reference protein EG1 and the protein 154E4 in the extracellular medium

10

The genes of the endoglucanase of *Trichoderma reesei* (EG1) and of the variant 154E4 were cloned without their signal peptide in the vector pESC-Leu Δ myc (CNRS-CERMAV). This construct allows the expression of the proteins in the culture medium of the strain *Saccharomyces cerevisiae* EBY100, which is auxotrophic for leucine and tryptophan (Boder ET and Wittrup KD, Biotechnol Prog, 1998, 14:55-62). This plasmid makes it possible to place the expression of the genes under the control of the galactose-inducible GAL1 promoter, and has the selection-marker gene for the auxotrophy (Leu2) that

15

allows the selection of the transformants.

The transformation of *Saccharomyces cerevisiae* EBY100 was carried out according to the conventional methods known to a person skilled in the art (transformation of yeasts via thermal shock and lithium acetate). The transformants were selected on a YNB 0.67%-Glc 2%-Trp 0.01% medium.

20

One transformant for each gene (Sc α -EG1 and Sc α -154E4) was used to inoculate 15mL of a YNB 0.67%-Glc 2%-SD-Trp 0.01% minimal medium. SD is a mixture of amino acids (40mg/L of adenine

sulphate; 20mg/L of L-arginine; 100mg/L of L-aspartic acid; 100mg/L of L-glutamic acid; 20mg/L of L-histidine; 30mg/L of L-lysine; 20mg/L of L-methionine; 50mg/L of L-phenylalanine; 375mg/L of L-serine; 200mg/L of L-threonine; 30mg/L of L-tyrosine; 150mg/L of L-valine and 20mg/L of uracil). After 24 hours of preculture at 30°C with stirring at 220rpm, the 2 strains Sca-EG1 and Sca-154E4 were used to inoculate (OD₆₀₀ of 0.5) 150mL of YNB 0.67%-Gal 2%-SD-Trp 0.01% medium. The cultures were incubated at 25°C with stirring at 220rpm. After 8 hours of incubation, 6mL of sodium citrate at pH 5.6 were added to each culture in order to stabilize the pH at 5.

After 4 days of incubation, 20mL of culture were removed. The culture supernatant was obtained after centrifugation at 3000g, at 4°C for 5 minutes.

2- Determination of the endoglucanase activity on p-nitrophenyl-β-lactoside

The endoglucanase activity of the culture supernatants was measured by hydrolysis of the substrate p-nitrophenyl-β-lactoside (pNPL) in a volume of 700μL under the following conditions:

- 50mM of citrate buffer at pH5;
- 2mM of pNPL;
- 605μL and 90μL of culture supernatant from the Sca-154E4 strains;
- Incubation at 35°C or 50°C for 30min for the strain Sca-EG1.

The reaction was stopped by adding 100μL of 1M sodium carbonate to 100μL of the reaction medium. The concentration of para-nitrophenol (pNP) released by hydrolysis of the pNPL was determined by measuring the absorbance at 415nm in comparison to a para-nitrophenol standard range (linear from 0.36μM to 360μM).

TABLE 8 presents the results for endoglucanase activity (EA in nmol.min⁻¹.mL⁻¹ of culture) on pNPL at 35°C and 50°C, for the culture media of the strains ScaEG1 and Sca154E4.

TABLE 8: Endoglucanase activity on pNPL at 35°C and 50°C, for the culture media of the strains ScaEG1 and Sca154E4

	EA at 35°C	EA at 50°C	Ratio of activity 35°C/50°C	Improvement in the EA at 35°C
Sca-EG1	1.17	1.18	1	-
Sca-154E4	45.0	61.1	1.4	38.5

The results obtained show an improvement, at 35°C, in the enzymatic activity of the strain Sca-154E4 by a factor of close to 40 with respect to the strain expressing the reference protein EG1 of *T. reesei* (SEQ ID NO:2). The magnitude of the improvement in activity observed with respect to *E. coli* and

T. reesei suggests that the enzyme not only has an improved specific activity, but that it is also overexpressed and/or better secreted.

3- Determination of the endoglucanase activity on carboxymethyl cellulose

The endoglucanase activity of the culture supernatants was measured via hydrolysis of carboxymethyl cellulose (CMC) in a volume of 700µL under the following conditions:

- 50mM of citrate buffer at pH 5;
- 1% CMC;
- 210µL of culture supernatant of the strains Sca-EG1 and Sca-154E4 respectively dialysed against pH 5 50mM citrate buffer on a 10kDa membrane and concentrated 2 times;
- Incubation at 35°C for 48 hours.

The reaction was stopped by adding 150µL of reagent containing DNS to 100µL of the reaction medium. After heating for 5 minutes at 100°C and cooling in ice, the quantity of reducing sugars released was determined by measuring the absorbance at 550nm in comparison to a standard range created with glucose.

FIGURE 1 presents the results of hydrolysis of the 1% CMC by the reference endoglucanase EG1 (SEQ ID NO:2) and its mutant 154E4 (SEQ ID NO:10) respectively secreted in the culture medium of the strains Sca-EG1 and Sca-154E4.

The results of FIGURE 1 show that during the first hours of reaction, the quantity of reducing sugars released per 1mL of culture of the strain Sca-154E4 is approximately 10 times greater than with 1mL of Sca-EG1. The magnitude of the improvement in activity observed with respect to *E. coli* and *T. reesei* suggests that the enzyme not only has an improved specific activity, but that it is also overexpressed and/or better secreted.

Example 5: Production of enzymes by *T. reesei* in fed flasks

The reference strains and those having the best activity on CMC (CL847, ΔEG1, ΔEG1cEG1, 154E4/2, 154E4/8, 11G8/10, 11G8/12, 11G8/13) were cultured in 250mL Erlenmeyer flasks. 55mL of F45 medium (10g/L of pH 6 dipotassium phthalate buffer, 4.2g/L (NH₄)₂SO₄, 300mg/L MgSO₄·7H₂O, 150mg/L CaCl₂·2H₂O, 1.5g/L cornsteep, 0.07% orthophosphoric acid, 5mg/L FeSO₄, 1.4mg/L MnSO₄, 1.4mg/L ZnSO₄, 3.7mg/L CoCl₂ and 12.5g/L glucose) are inoculated and stirred at 150rpm and 30°C. The production of enzymes occurs in two phases: a batch phase on glucose and a fed-batch phase on lactose. Regular samples allow the moment at which the concentration of glucose goes below 3g/L to be determined. At this stage, fed-batch feeding using a syringe driver (6 channels) is initiated. The cultures are fed with a solution of 50g/L lactose and 0.3% NH₃ at a flow rate of 40mg of sugar/g of biomass per hour. Daily samples are taken in order to determine the pH, the dry weight and the concentration of

proteins in the supernatant. After 5 days of fed-batch culture, the culture is filtered over a 0.45µm filter and the supernatant is frozen.

The final concentration of proteins was approximately 3 to 4g/L. If the concentration was below 3g/L, the supernatants were concentrated in a column (Vivaspin MWCO5, Sartorius).

5

Example 6: Effectiveness of the enzymes resulting from the L-shuffling in hydrolysis of lignocellulosic biomass according to an SHF method

The reference substrate used is a wheat straw that has undergone a pretreatment via steam explosion (19 bar - 3 minutes) after acid impregnation with 0.01% H₂SO₄ for 10 hours, washed, neutralised at pH 5, pressed, dried. Table 9 presents the composition of the reference substrate.

10

Composition	% m / m
WIS	97.52
Ash content	5
Cellulose	51.7
Corrected xylans	3.57
Hemicellulose	4.14
Klason lignin (overestimated)	36.49
Acetyl	0.6

Table 9: Composition of the straw used for the hydrolysis tests

The hydrolyses were carried out at 10% m/m dry matter, or an equivalent of 5.4% m/m cellulose. The concentration of proteins is fixed at 10mg/g of dry matter or approximately 19mg/g of cellulose. The concentration of the enzyme cocktails was measured via the Lowry method, using BSA as a reference. Each cocktail was supplemented in terms of β-glucosidase activity with 120 ± 2 IU/g of cellulose, by adding β-glucosidase SP188 (Novozymes).

15

The tests are carried out in Eppendorf tubes having 2mL of useful volume (1g reaction volume) and containing:

20

- 0.11 ± 0.001g of washed straw substrate
- 0.9 ± 0.02mL of hydrolysis reaction medium composed of 50mM – pH 4.8 acetate buffer and chloramphenicol (0.05g/L)
- between 0.1 and 0.2 ± 0.02g of enzyme cocktail according to their protein concentration.

25

The enzymatic hydrolyses are carried out at 45 ± 2°C with vortex stirring at 900 revolutions per minute in an *Eppendorf Thermomixer Comfort*.

All the tests are carried out in duplicate with sampling times fixed at t 24, 48 and 96 hours, with sampling at t 72 hours for some.

At each sampling time, the hydrolysates are scalded for 5 minutes in the sacrificed Eppendorf tubes. These tubes are then cooled and centrifuged. The assay of the glucose is carried out via HPLC. In parallel, the solid residues from each Eppendorf tube are washed and centrifuged 3 times before being dried at 105°C for 24 hours in such a way as to evaluate the WIS (Water Insoluble Solids). The calculation of the hydrolysis yield is carried out while taking into account the WIS.

The cocktails from example 5 were evaluated. Two control tests are carried out with the reference cocktails also supplemented with β -glucosidase for comparison: a cocktail produced by the strain CL847 Δ EG1 (Δ EG1) and a cocktail produced by the strain CL847 Δ EG1 retransformed with the reference gene EG1 (Δ EG1cEG1).

FIGURE 2 presents the results of SHF for the cocktail resulting from the strain 154/8 expressing the variant 154E4 (SEQ ID NO:10).

The results presented in FIGURE 2 show that the initial rate of hydrolysis of the cocktail produced by the variant 154E4 is greater than those of the reference cocktails Δ EG1 and Δ EG1cEG1. The final hydrolysis yield is also greater than that of the reference cocktails Δ EG1 and Δ EG1cEG1.

FIGURE 3 presents the results of SHF for the cocktail resulting from the strain 11G8/10 expressing the variant 11G8 (SEQ ID NO:22).

The results presented in FIGURE 3 show that the initial rate of hydrolysis of the cocktail produced by the variant 11G8 is greater than those of the reference cocktails Δ EG1 and Δ EG1cEG1. The final hydrolysis yield is also greater than that of the reference cocktails Δ EG1 and Δ EG1cEG1.

Example 7: Efficiency of the enzymes in hydrolysis of lignocellulosic biomass according to an SSF method

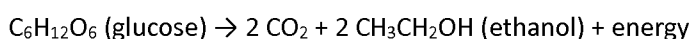
The substrate used is the same as that described in table 9 (Example 6).

The SSFs are carried out in triplicate in laboratory reactors. Said reactors consist of the following elements:

- a glass flask having a useful volume of 30mL;
- a safety cap made of polyether ether ketone (Peek);
- a DV-118 one-way valve marketed by the company Vaplock fastened through the cap. The valve is configured to open as an outlet when the relative pressure in the flask is greater than 70mbar;
- a hollow tube made of polypropylene, fitted through a second one passing through the cap and provided with a septum at the lower end of said tube;
- a flat seal positioned between the neck of the flask and the cap.

The principle of implementation of the bioreactors is the following: the CO₂ produced during the ethanolic fermentation accumulates in the vapour space located above the reaction medium

leading, via accumulation, to an increase in the pressure in the bioreactor (P_G). When P_G exceeds the cracking pressure of the one-way valve (P_S), said valve opens in order to allow a quantity of gas to escape that is for example determined by weighing. When $P_G < P_S$, the valve closes until P_G is greater than P_S . Thus, during operation, the bioreactor is always pressurised in such a way as to ensure a stable anaerobic medium for the fermentation. The quantity of ethanol produced is evaluated by the production of CO_2 estimated via loss of weight using the following stoichiometric equation for fermentation of glucose into ethanol:



The culture medium used for the SSF is an aqueous medium that comprises:

- a 50mM acetate buffer for pH 5;
- chloramphenicol at 0.1g/L;
- nutritive medium with 3g/L of KH_2PO_4 , 2g/L of $(NH_4)_2SO_4$, 0.4g/L of $MgSO_4 \cdot 7H_2O$ and 1g/L of yeast extract.

The SSFs were carried out at $10 \pm 0.01\%$ m/m of dry matter, or an equivalent of 5.4% cellulose m/m for a total reaction mass of 15 ± 0.003 g. The concentration of proteins is fixed at 10 ± 0.01 mg of cellulases per gram of dry matter, or approximately 19mg/g cellulose. The concentration of the enzyme cocktails was measured via the Lowry method while using BSA (Bovine Serum Albumin) as a reference. Each cocktail was supplemented in terms of β -glucosidase activity with 120 ± 2 IU/g cellulose, by adding β -glucosidase SP188 (Novozymes).

The yeast for fermentation of the sugars (*Saccharomyces cerevisiae*, Ethanol Red strain, Fermentis, France) is added to the medium in order to obtain a concentration of 2 ± 0.1 g/kg.

The enzymes and the yeasts are added into the bioreactors after one hour of conditioning the pretreated wheat straw at 35°C with the buffer, the chloramphenicol and the culture medium.

The SSF reaction is carried out at a temperature of approximately 35°C, by placing the laboratory bioreactor into an INFORS Multitron HT Standard incubator with an orbital rotation speed of 150 revolutions per minute.

Over time, the mass loss was monitored by weighing the bioreactors. At the end of the reaction, the fermentation must be heated at 100°C for 5 minutes, cooled and centrifuged in order to separate the non-hydrolysed solids from the fermentation liquid. The latter is then analysed via gas chromatography in order to determine its ethanol concentration.

The cocktails from example 5 were evaluated. Two control tests are carried out with the reference cocktails also supplemented with β -glucosidase for comparison: a cocktail produced by the

strain CL847 Δ EG1 (Δ EG1) and a cocktail produced by the strain CL847 Δ EG1 retransformed with the reference gene EG1 (Δ EG1cEG1).

FIGURE 4 presents the results of SSF for the 2 cocktails expressing the endoglucanase 154E4 (average of the results obtained with the 2 variants).

5 The results presented in FIGURE 4 show that the degree of conversion (ethanol production for the same dose of enzymes) of the SSF over 100 hours for the 2 cocktails expressing the endoglucanase 154E4 is greater than those of the reference cocktails Δ EG1 and Δ EG1cEG1.

FIGURE 5 presents the results of SSF for the 3 cocktails expressing the endoglucanase 11G8 (average average of the results obtained with the 2 variants).

10 The results presented in FIGURE 5 show that the degree of conversion of the SSF over 100 hours for the 3 cocktails expressing the endoglucanase 11G8 (average) is greater than those of the reference cocktails Δ EG1 and Δ EG1cEG1.

SEQUENCE LISTING

<110> IFPEN
 5 <120> ENDOGLUCANASE VARIANTS HAVING IMPROVED ACTIVITY, AND USES OF SAME
 <130> BFF
 <160> 30
 10 <170> PatentIn version 3.5
 <210> 1
 <211> 1380
 15 <212> DNA
 <213> *Trichoderma reesei*
 <400> 1
 20 atggcgccct cagttacact gccgttgacc acggccatcc tggccattgc ccggctcgtc
 60
 gccgcccagc aaccgggtac cagcaccccc gaggtccatc ccaagttgac aacctacaag
 120
 25 tgtacaaagt ccgggggggtg cgtggcccag gacacctcgg tggtccttga ctggaactac
 180
 cgctggatgc acgacgcaaa ctacaactcg tgcaccgtca acggcggcgt caacaccacg
 240
 30 ctctgccctg acgaggcgac ctgtggcaag aactgcttca tcgagggcgt cgactacgcc
 300
 gcctcgggcg tcacgacctc gggcagcagc ctcaccatga accagtacat gcccagcagc
 35 360
 tctggcggct acagcagcgt ctctcctcgg ctgtatctcc tggactctga cggtgagtac
 420
 40 gtgatgctga agctcaacgg ccaggagctg agcttcgacg tcgacctctc tgctctgccg
 480
 tgtggagaga acggctcgtc ctacctgtct cagatggacg agaacggggg cgccaaccag
 540
 45 tataacacgg ccggtgcaa ctacgggagc ggctactgcg atgctcagtg ccccgctccag
 600
 acatggagga acggcaccct caacactagc caccagggct tctgctgcaa cgagatggat
 50 660
 atcctggagg gcaactcgag ggcgaaatgcc ttgaccctc actcttgcac ggccacgggc
 720
 55 tgcgactctg ccggttgccg ctccaacccc tatggcagcg gctacaaaag ctactacggc
 780
 cccggagata ccgttgacac ctccaagacc ttcaccatca tcaccagtt caacacggac
 840

aacggctcgc cctcgggcaa ccttgtgagc atcaccgcga agtaccagca aaacggcgtc
 900

5 gacatcccca gcgcccagcc cggcggcgac accatctcgt cctgcccgtc cgcctcagcc
 960

tacggcggcc tcgccacat gggcaaggcc ctgagcagcg gcatggtgct cgtgttcagc
 1020

10 atttgaacg acaacagcca gtacatgaac tggctcgaca gcggcaacgc cggcccctgc
 1080

agcagcaccg agggcaaccc atccaacatc ctggccaaca accccaacac gcacgtcgtc
 15 1140

ttctccaaca tccgctgggg agacattggg tctactacga actcgactgc gccccgccc
 1200

20 ccgcctgcgt ccagcagcag gttttcgact acacggagga gctcgacgac ttcgagcagc
 1260

ccgagctgca cgcagactca ctgggggcag tgcggtggca ttgggtacag cgggtgcaag
 1320

25 acgtgcacgt cgggcactac gtgccagtat agcaacgact actactcgca atgcctttaa
 1380

30 <210> 2
 <211> 459
 <212> PRT
 <213> Trichoderma reesei

35 <400> 2

Met Ala Pro Ser Val Thr Leu Pro Leu Thr Thr Ala Ile Leu Ala Ile
 1 5 10 15

40 Ala Arg Leu Val Ala Ala Gln Gln Pro Gly Thr Ser Thr Pro Glu Val
 20 25 30

45 His Pro Lys Leu Thr Thr Tyr Lys Cys Thr Lys Ser Gly Gly Cys Val
 35 40 45

50 Ala Gln Asp Thr Ser Val Val Leu Asp Trp Asn Tyr Arg Trp Met His
 50 55 60

Asp Ala Asn Tyr Asn Ser Cys Thr Val Asn Gly Gly Val Asn Thr Thr
 65 70 75 80

55 Leu Cys Pro Asp Glu Ala Thr Cys Gly Lys Asn Cys Phe Ile Glu Gly
 85 90 95

	Val	Asp	Tyr	Ala	Ala	Ser	Gly	Val	Thr	Thr	Ser	Gly	Ser	Ser	Leu	Thr	
				100					105					110			
5	Met	Asn	Gln	Tyr	Met	Pro	Ser	Ser	Ser	Gly	Gly	Tyr	Ser	Ser	Val	Ser	
			115					120					125				
10	Pro	Arg	Leu	Tyr	Leu	Leu	Asp	Ser	Asp	Gly	Glu	Tyr	Val	Met	Leu	Lys	
		130					135					140					
15	Leu	Asn	Gly	Gln	Glu	Leu	Ser	Phe	Asp	Val	Asp	Leu	Ser	Ala	Leu	Pro	
	145					150					155					160	
20	Cys	Gly	Glu	Asn	Gly	Ser	Leu	Tyr	Leu	Ser	Gln	Met	Asp	Glu	Asn	Gly	
					165					170					175		
25	Gly	Ala	Asn	Gln	Tyr	Asn	Thr	Ala	Gly	Ala	Asn	Tyr	Gly	Ser	Gly	Tyr	
				180					185					190			
30	Cys	Asp	Ala	Gln	Cys	Pro	Val	Gln	Thr	Trp	Arg	Asn	Gly	Thr	Leu	Asn	
			195					200					205				
35	Thr	Ser	His	Gln	Gly	Phe	Cys	Cys	Asn	Glu	Met	Asp	Ile	Leu	Glu	Gly	
		210					215					220					
40	Asn	Ser	Arg	Ala	Asn	Ala	Leu	Thr	Pro	His	Ser	Cys	Thr	Ala	Thr	Ala	
	225					230					235					240	
45	Cys	Asp	Ser	Ala	Gly	Cys	Gly	Phe	Asn	Pro	Tyr	Gly	Ser	Gly	Tyr	Lys	
					245					250					255		
50	Ser	Tyr	Tyr	Gly	Pro	Gly	Asp	Thr	Val	Asp	Thr	Ser	Lys	Thr	Phe	Thr	
				260					265					270			
55	Ile	Ile	Thr	Gln	Phe	Asn	Thr	Asp	Asn	Gly	Ser	Pro	Ser	Gly	Asn	Leu	
			275					280					285				
60	Val	Ser	Ile	Thr	Arg	Lys	Tyr	Gln	Gln	Asn	Gly	Val	Asp	Ile	Pro	Ser	
		290					295					300					
65	Ala	Gln	Pro	Gly	Gly	Asp	Thr	Ile	Ser	Ser	Cys	Pro	Ser	Ala	Ser	Ala	
	305					310					315					320	
70	Tyr	Gly	Gly	Leu	Ala	Thr	Met	Gly	Lys	Ala	Leu	Ser	Ser	Gly	Met	Val	
					325					330					335		

5 Leu Val Phe Ser Ile Trp Asn Asp Asn Ser Gln Tyr Met Asn Trp Leu
 340 345 350

10 Asp Ser Gly Asn Ala Gly Pro Cys Ser Ser Thr Glu Gly Asn Pro Ser
 355 360 365

15 Asn Ile Leu Ala Asn Asn Pro Asn Thr His Val Val Phe Ser Asn Ile
 370 375 380

20 Arg Trp Gly Asp Ile Gly Ser Thr Thr Asn Ser Thr Ala Pro Pro Pro
 385 390 395 400

25 Pro Pro Ala Ser Ser Thr Thr Phe Ser Thr Thr Arg Arg Ser Ser Thr
 405 410 415

30 Thr Ser Ser Ser Pro Ser Cys Thr Gln Thr His Trp Gly Gln Cys Gly
 420 425 430

35 Gly Ile Gly Tyr Ser Gly Cys Lys Thr Cys Thr Ser Gly Thr Thr Cys
 435 440 445

40 Gln Tyr Ser Asn Asp Tyr Tyr Ser Gln Cys Leu
 450 455

45 <210> 3
 <211> 1329
 <212> DNA
 <213> Artificial Sequence

50 <220>
 <223> 76B4

55 <400> 3
 cagcaaccgg gttcgggtgac ccccgaggtc catcccaagt tgcccacctg gcagtgtaca
 60
 aagtccgggg ggtgcggtgga gcaggacacc tcgggtggtcc ttgactggaa ctaccgctgg
 120
 ttccacacct cggacaacac cacctcgtgc accacgtcct cgggcatcga ctcgacgctc
 180
 tgccctgacg cagcgacctg tgccgagaac tgcgtcgtgg agggcaccga ctacaccagc
 240
 tcgggcatcg agacctcggg cagcagcctc accctgaggc agttcgtcaa ggacagcgag
 300

ggcaagacca acagcgtctc tcctcgggcc tatctcctgg gcgcgcacgg tgactacgtg
 360
 5 atgttcaagc tcctcaacca ggagctgagc ttcgacgtcg acgtctctac gctgccgtgt
 420
 ggagagaacg cagcgtctta cttctctgag atggacaaga ccggggggccg gaacgagcac
 480
 10 aacacggggc gtgccaagta cgggagcggc tactgcgatg ctcaagtccc cgtccagaca
 540
 tggaacaacg gcaccctcaa cactagccac cagggctcgt gctgcaacga gatggatatc
 600
 15 ctggaggcca actcgagggc gaattcgtac acccctcacc cctgcagtgc cacggactgc
 660
 gacaagggcg gttgcggcct caaccctat gctctcggcc aaaaaagcta ctggggcccc
 720
 20 ggaggcaccg ttgacacctc caagccctc accatcacca ccagttcat cacgaacgac
 780
 25 ggcaccacca ccggcaccct ttccgaaatc cgccgccagt acatgcaaaa cggcaagggtg
 840
 atcgccaatg ccgtttcctc cactggcgtc aactccatca ccgaggactg gtgcacgtcc
 900
 30 gtcgacggct cggccgccac ctttggcggc ctcaccacca tgggcaaggc cctggggccg
 960
 ggcatggtgc tcatcttcag catttggaac gacgccagcg gctttatgaa ctggctcgac
 1020
 35 agcggcaacg ccggcccctg cagcagcacc gagggcaacc cagacttgat caaggcccag
 1080
 40 aaccccacga cgcacgtcgt cttctccaac atccgctggg gagacattgg gtctactttc
 1140
 aagggttctg atggctcggg gacgacgacg acgtcgacta catcgaccaa gaccacgact
 1200
 45 tcgaccgcgc cggggccaac gcagactcac tatgggcagt gcggtggcca aggggtggact
 1260
 gggcccacgg cttgcgcac cccctacacg tgccagggtc tgaaccggtg gtactcgcaa
 1320
 50 tgccttttaa
 1329
 55
 <210> 4
 <211> 442
 <212> PRT
 <213> Artificial Sequence

<220>

<223> 76B4

5 <400> 4

Gln Gln Pro Gly Ser Val Thr Pro Glu Val His Pro Lys Leu Pro Thr
 1 5 10 15

10 Trp Gln Cys Thr Lys Ser Gly Gly Cys Val Glu Gln Asp Thr Ser Val
 20 25 30

15 Val Leu Asp Trp Asn Tyr Arg Trp Phe His Thr Ser Asp Asn Thr Thr
 35 40 45

20 Ser Cys Thr Thr Ser Ser Gly Ile Asp Ser Thr Leu Cys Pro Asp Ala
 50 55 60

25 Ala Thr Cys Ala Glu Asn Cys Val Val Glu Gly Thr Asp Tyr Thr Ser
 65 70 75 80

Ser Gly Ile Glu Thr Ser Gly Ser Ser Leu Thr Leu Arg Gln Phe Val
 85 90 95

30 Lys Asp Ser Glu Gly Lys Thr Asn Ser Val Ser Pro Arg Ala Tyr Leu
 100 105 110

35 Leu Gly Ala Asp Gly Asp Tyr Val Met Phe Lys Leu Leu Asn Gln Glu
 115 120 125

40 Leu Ser Phe Asp Val Asp Val Ser Thr Leu Pro Cys Gly Glu Asn Ala
 130 135 140

45 Ala Leu Tyr Phe Ser Glu Met Asp Lys Thr Gly Gly Arg Asn Glu His
 145 150 155 160

Asn Thr Gly Gly Ala Lys Tyr Gly Ser Gly Tyr Cys Asp Ala Gln Cys
 165 170 175

50 Pro Val Gln Thr Trp Asn Asn Gly Thr Leu Asn Thr Ser His Gln Gly
 180 185 190

55 Ser Cys Cys Asn Glu Met Asp Ile Leu Glu Ala Asn Ser Arg Ala Asn
 195 200 205

Ser Tyr Thr Pro His Pro Cys Ser Ala Thr Asp Cys Asp Lys Gly Gly

		210					215					220					
5	Cys 225	Gly	Phe	Asn	Pro	Tyr 230	Ala	Leu	Gly	Gln	Lys 235	Ser	Tyr	Trp	Gly	Pro 240	
10	Gly	Gly	Thr	Val	Asp 245	Thr	Ser	Lys	Pro	Phe 250	Thr	Ile	Thr	Thr	Gln 255	Phe	
15	Ile	Thr	Asn	Asp 260	Gly	Thr	Thr	Thr	Gly 265	Thr	Leu	Ser	Glu	Ile 270	Arg	Arg	
20	Gln	Tyr	Met 275	Gln	Asn	Gly	Lys	Val 280	Ile	Ala	Asn	Ala	Val 285	Ser	Ser	Thr	
25	Gly	Val 290	Asn	Ser	Ile	Thr	Glu 295	Asp	Trp	Cys	Thr	Ser 300	Val	Asp	Gly	Ser	
30	Ala 305	Ala	Thr	Phe	Gly	Gly 310	Leu	Thr	Thr	Met	Gly 315	Lys	Ala	Leu	Gly	Arg 320	
35	Gly	Met	Val	Leu	Ile 325	Phe	Ser	Ile	Trp	Asn 330	Asp	Ala	Ser	Gly	Phe 335	Met	
40	Asn	Trp	Leu	Asp 340	Ser	Gly	Asn	Ala	Gly 345	Pro	Cys	Ser	Ser	Thr 350	Glu	Gly	
45	Asn	Pro	Asp 355	Leu	Ile	Lys	Ala	Gln 360	Asn	Pro	Thr	Thr	His 365	Val	Val	Phe	
50	Ser	Asn 370	Ile	Arg	Trp	Gly	Asp 375	Ile	Gly	Ser	Thr	Phe 380	Lys	Gly	Ser	Asp	
55	Gly 385	Ser	Val	Thr	Thr	Thr 390	Thr	Ser	Thr	Thr	Ser 395	Thr	Lys	Thr	Thr	Thr 400	
60	Ser	Thr	Ala	Pro	Gly 405	Pro	Thr	Gln	Thr	His 410	Tyr	Gly	Gln	Cys	Gly 415	Gly	
65	Gln	Gly	Trp	Thr 420	Gly	Pro	Thr	Ala	Cys 425	Ala	Ser	Pro	Tyr	Thr 430	Cys	Gln	
70	Val	Leu	Asn 435	Pro	Trp	Tyr	Ser	Gln 440	Cys	Leu							

<210> 5
 <211> 1338
 <212> DNA
 <213> Artificial Sequence
 5
 <220>
 <223> 105F11
 <400> 5
 10 cagcaaccgg gtaccagcac ccccgaggtc catcccaagt tgacaaccta caagtgtaca
 60
 aagtccgggg ggtgctgtggc ccaggacacc tcggtgggtcc ttgactggaa ctaccactgg
 120
 15 attcacacgg tcgatgggta cacatcgtgc accacatcgt ccggcgctga cagcacgctc
 180
 tgccctgacg cggcgacctg tgccaagaac tgctgtatcg agccggccaa ctacaccagc
 20 240
 gccggcgctca cgacctcggg cgacagcctc accatgtacc agtacgttca gagcaacggc
 300
 25 gtctacacca acgcctctcc tcggctgtat ctctggggcc ccgacaagaa ctacgtgatg
 360
 ctgaagctcc taggccagga gctgaccttc gacgtcgacc tctctgctct gccgtgtgga
 420
 30 gagaacggct cgctctacct gtctcagatg gacgagaacg ggggcgccaa ccagtataac
 480
 acggccggtg ccaactacgg gagcggctac tgcgatgctc agtgccccgt ccagacatgg
 35 540
 aggaacggca ccctcaacac tagccaccag ggctcgtgct gcaacgagat ggatatcctg
 600
 40 gaggccaaact cgaaggcgga agccttcacc cctcaccctt gcatcggcga caactgcgac
 660
 aaggcggtt gcggcttcaa cccctatgct ctgggcaaaa aaagctactg gggccccgga
 720
 45 ggcaccgttg acacctcaa gcccttcacc atcaccaccc agttcatcac gaacgacggc
 780
 accaccaccg gcaccctttc cgaaatccgc cgccagtaca tgcaaaacgg caaggtgatc
 50 840
 gccaatgccg tttcctccac tggcgtcaac tccatcaccg aggactgggtg caggtccgtc
 900
 55 gacggctcgg ccgccacctt tggcggcctc accaccatgg gcaaggccct gggccgcggc
 960
 atggtgctcg tgttcagcat ttggaacgac aacagccagt acatgaactg gctcgacagc
 1020

ggcaacgccg gccctgcag cagcaccgag ggcaacccat ccaacatcct ggccaacaac
1080

5 cccaacacgc acgtcgtctt ctccaacatc cgctggggag acattgggtc tactacgaac
1140

tcgactgcgc ccccgcccc gcctgcgtcc agcacgacgt tttcgactac acggaggagc
1200

10 tcgacgactt cgagcagccc gagctgcacg cagactcact gggggcagtg cggtggcatt
1260

gggtacagcg ggtgcaagac gtgcacgtcg ggcaactacgt gccagtatat caacgactac
1320

15 tactcgcaat gcctttaa
1338

20

<210> 6
<211> 445
<212> PRT
<213> Artificial Sequence

25

<220>
<223> 105F11

<400> 6

30

Gln	Gln	Pro	Gly	Thr	Ser	Thr	Pro	Glu	Val	His	Pro	Lys	Leu	Thr	Thr
1				5					10					15	

35

Tyr	Lys	Cys	Thr	Lys	Ser	Gly	Gly	Cys	Val	Ala	Gln	Asp	Thr	Ser	Val
			20					25					30		

40

Val	Leu	Asp	Trp	Asn	Tyr	His	Trp	Ile	His	Thr	Val	Asp	Gly	Tyr	Thr
		35					40					45			

45

Ser	Cys	Thr	Thr	Ser	Ser	Gly	Val	Asp	Ser	Thr	Leu	Cys	Pro	Asp	Ala
	50					55					60				

50

Ala	Thr	Cys	Ala	Lys	Asn	Cys	Val	Ile	Glu	Pro	Ala	Asn	Tyr	Thr	Ser
65					70					75					80

55

Ala	Gly	Val	Thr	Thr	Ser	Gly	Asp	Ser	Leu	Thr	Met	Tyr	Gln	Tyr	Val
				85					90					95	

60

Gln	Ser	Asn	Gly	Val	Tyr	Thr	Asn	Ala	Ser	Pro	Arg	Leu	Tyr	Leu	Leu
			100					105					110		

65

Gly	Pro	Asp	Lys	Asn	Tyr	Val	Met	Leu	Lys	Leu	Leu	Gly	Gln	Glu	Leu
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

				115				120					125				
5	Thr	Phe	Asp	Val	Asp	Leu	Ser	Ala	Leu	Pro	Cys	Gly	Glu	Asn	Gly	Ser	
	130						135					140					
10	Leu	Tyr	Leu	Ser	Gln	Met	Asp	Glu	Asn	Gly	Gly	Ala	Asn	Gln	Tyr	Asn	
	145					150					155					160	
15	Thr	Ala	Gly	Ala	Asn	Tyr	Gly	Ser	Gly	Tyr	Cys	Asp	Ala	Gln	Cys	Pro	
					165					170					175		
20	Val	Gln	Thr	Trp	Arg	Asn	Gly	Thr	Leu	Asn	Thr	Ser	His	Gln	Gly	Ser	
				180					185					190			
25	Cys	Cys	Asn	Glu	Met	Asp	Ile	Leu	Glu	Ala	Asn	Ser	Lys	Ala	Glu	Ala	
			195					200					205				
30	Phe	Thr	Pro	His	Pro	Cys	Ile	Gly	Asp	Asn	Cys	Asp	Lys	Gly	Gly	Cys	
	210						215					220					
35	Gly	Phe	Asn	Pro	Tyr	Ala	Leu	Gly	Gln	Lys	Ser	Tyr	Trp	Gly	Pro	Gly	
	225					230					235					240	
40	Gly	Thr	Val	Asp	Thr	Ser	Lys	Pro	Phe	Thr	Ile	Thr	Thr	Gln	Phe	Ile	
					245					250				255			
45	Thr	Asn	Asp	Gly	Thr	Thr	Thr	Gly	Thr	Leu	Ser	Glu	Ile	Arg	Arg	Gln	
				260				265						270			
50	Tyr	Met	Gln	Asn	Gly	Lys	Val	Ile	Ala	Asn	Ala	Val	Ser	Ser	Thr	Gly	
			275					280					285				
55	Val	Asn	Ser	Ile	Thr	Glu	Asp	Trp	Cys	Thr	Ser	Val	Asp	Gly	Ser	Ala	
	290						295					300					
60	Ala	Thr	Phe	Gly	Gly	Leu	Thr	Thr	Met	Gly	Lys	Ala	Leu	Gly	Arg	Gly	
	305					310					315					320	
65	Met	Val	Leu	Val	Phe	Ser	Ile	Trp	Asn	Asp	Asn	Ser	Gln	Tyr	Met	Asn	
					325					330					335		
70	Trp	Leu	Asp	Ser	Gly	Asn	Ala	Gly	Pro	Cys	Ser	Ser	Thr	Glu	Gly	Asn	
				340				345						350			

Pro Ser Asn Ile Leu Ala Asn Asn Pro Asn Thr His Val Val Phe Ser
355 360 365

5 Asn Ile Arg Trp Gly Asp Ile Gly Ser Thr Thr Asn Ser Thr Ala Pro
370 375 380

10 Pro Pro Pro Pro Ala Ser Ser Thr Thr Phe Ser Thr Thr Arg Arg Ser
385 390 395 400

15 Ser Thr Thr Ser Ser Ser Pro Ser Cys Thr Gln Thr His Trp Gly Gln
405 410 415

20 Cys Gly Gly Ile Gly Tyr Ser Gly Cys Lys Thr Cys Thr Ser Gly Thr
420 425 430

Thr Cys Gln Tyr Ser Asn Asp Tyr Tyr Ser Gln Cys Leu
435 440 445

25 <210> 7
<211> 1131
<212> DNA
<213> Artificial Sequence

30 <220>
<223> 107H12

35 <400> 7
cagcaaccgg gttcgggtgac ccccgaggtc catcccaagt tgacaaccta caagtgtaca
60

aagtccgggg ggtgcgtggc ccaggacacc tcggtggtcc ttgactgggg ctaccactgg
120

40 attcacacgg tcgatgggta cacatcgtgc accacatcgt ccggcggtcga cagcacgctc
180

45 tgccttgacg cggcgacctg tgcaagaac tgcgatgatc agccggccaa ctacaccagc
240

gccggcggtca cgacctcggg cgacagcctc accctgaggg agttcgtcaa ggacagcgag
300

50 ggcaagacca acagcgtctc tcctcggggc tatctcctgg gcgccgacgg tgactacgtg
360

atgttcaagc tcctcaacca ggagctgagc ttcgacgtcg acctctctac gctgcggtgt
420

55 ggagagaacg cagcgtctta cttctctgag atggacaaga ccggggggccg gaacgagcac
480

aacacggggc gtgccaagta cgggagcggc tactgcatg ctcaagtccc cgtccagaca
540

tggaacaacg gcaccctcaa cactagccac cagggctcgt gctgcaacga gatggatata
 600
 5 ctggaggcca actcgaaggc ggaagccttc acccctcacc cctgcatcgg cgacaactgc
 660
 gacaagggcg gttgcggctt caactcgtat gcgcgcggca aaaaaaacta ctgggcgccc
 720
 10 ggaggcacc ctcgacacctc caagcccttc accatggtga cccagttcat cacggacgac
 780
 15 ggaccagct cgggcaagct tagccagatc gtgcgctcct acgtgcaaaa cggccagaag
 840
 gtcgccagcg ccgtgtccgg cggcgacagc atcacgctcg agggctgctc gtcctccgac
 900
 20 gcctacggcg gcctcgtcgg tatgggagcg gccctggggc gcggcatggt gctcgccatg
 960
 agcatttgga acgacgccag cggcttcatg aactggctcg acagcggcga caacggcccc
 1020
 25 tgcaacgaga ccgagggcga cccagccaac atcctggcca accacccoga ttgcaggctc
 1080
 30 gtcctgtcca acatccgctg gggagacatt gactctactg ttcagctcta a
 1131
 <210> 8
 <211> 376
 35 <212> PRT
 <213> Artificial Sequence
 <220>
 <223> 107H12
 40 <400> 8
 Gln Gln Pro Gly Ser Val Thr Pro Glu Val His Pro Lys Leu Thr Thr
 1 5 10 15
 45 Tyr Lys Cys Thr Lys Ser Gly Gly Cys Val Ala Gln Asp Thr Ser Val
 20 25 30
 50 Val Leu Asp Trp Gly Tyr His Trp Ile His Thr Val Asp Gly Tyr Thr
 35 40 45
 55 Ser Cys Thr Thr Ser Ser Gly Val Asp Ser Thr Leu Cys Pro Asp Ala
 50 55 60
 Ala Thr Cys Ala Lys Asn Cys Val Ile Glu Pro Ala Asn Tyr Thr Ser

	65					70						75				80
5	Ala	Gly	Val	Thr	Thr	Ser	Gly	Asp	Ser	Leu	Thr	Leu	Arg	Gln	Phe	Val
					85					90					95	
10	Lys	Asp	Ser	Glu	Gly	Lys	Thr	Asn	Ser	Val	Ser	Pro	Arg	Ala	Tyr	Leu
				100					105					110		
15	Leu	Gly	Ala	Asp	Gly	Asp	Tyr	Val	Met	Phe	Lys	Leu	Leu	Asn	Gln	Glu
			115					120					125			
20	Leu	Ser	Phe	Asp	Val	Asp	Leu	Ser	Thr	Leu	Pro	Cys	Gly	Glu	Asn	Ala
		130					135					140				
25	Ala	Leu	Tyr	Phe	Ser	Glu	Met	Asp	Lys	Thr	Gly	Gly	Arg	Asn	Glu	His
	145					150					155					160
30	Asn	Thr	Gly	Gly	Ala	Lys	Tyr	Gly	Ser	Gly	Tyr	Cys	Asp	Ala	Gln	Cys
					165					170					175	
35	Pro	Val	Gln	Thr	Trp	Asn	Asn	Gly	Thr	Leu	Asn	Thr	Ser	His	Gln	Gly
				180					185					190		
40	Ser	Cys	Cys	Asn	Glu	Met	Asp	Ile	Leu	Glu	Ala	Asn	Ser	Lys	Ala	Glu
			195					200					205			
45	Ala	Phe	Thr	Pro	His	Pro	Cys	Ile	Gly	Asp	Asn	Cys	Asp	Lys	Gly	Gly
		210					215					220				
50	Cys	Gly	Phe	Asn	Ser	Tyr	Ala	Arg	Gly	Asn	Lys	Asn	Tyr	Trp	Ala	Pro
	225					230					235					240
55	Gly	Gly	Thr	Leu	Asp	Thr	Ser	Lys	Pro	Phe	Thr	Met	Val	Thr	Gln	Phe
					245					250					255	
60	Ile	Thr	Asp	Asp	Gly	Thr	Ser	Ser	Gly	Lys	Leu	Ser	Gln	Ile	Val	Arg
				260					265					270		
65	Ser	Tyr	Val	Gln	Asn	Gly	Gln	Lys	Val	Ala	Ser	Ala	Val	Ser	Gly	Gly
			275					280					285			
70	Asp	Ser	Ile	Thr	Val	Glu	Gly	Cys	Ser	Ser	Ser	Asp	Ala	Tyr	Gly	Gly
		290					295					300				

Leu Val Gly Met Gly Glu Ala Leu Gly Arg Gly Met Val Leu Ala Met
 305 310 315 320

5 Ser Ile Trp Asn Asp Ala Ser Gly Phe Met Asn Trp Leu Asp Ser Gly
 325 330 335

10 Asp Asn Gly Pro Cys Asn Glu Thr Glu Gly Asp Pro Ala Asn Ile Leu
 340 345 350

15 Ala Asn His Pro Asp Ser Gln Val Val Leu Ser Asn Ile Arg Trp Gly
 355 360 365

Asp Ile Asp Ser Thr Val Gln Leu
 370 375

20 <210> 9
 <211> 1320
 <212> DNA
 <213> Artificial Sequence

25 <220>
 <223> 154E4

30 <400> 9
 cagcaaccgg gtaccagcac ccccgaggtc catcccaagt tgacaaccta caagtgtaca
 60

35 actgctgggg ggtgctggc ccaggacacc tcggtgggtcc ttgactgggg ctaccactgg
 120

attcacacgg tcgatgggta cacatcgtgc accacatcgt cgggcgtcga cagcacgctc
 180

40 tgccctgacg cggcgacctg tgcaagaac tgcgtgatcg agccggccaa ctacaccagc
 240

gccggcgtca cgacctcggg cgacagcctc accatgtacc agtacgttca gagcaacggc
 300

45 gtctacacca acgcctctcc tcgggcctat ctctggggc cgcacggtga ctacgtgatg
 360

ttcaagctcc tcaaccagga gctgagcttc gacgtcgacg tctctacgct gccgtgtgga
 420

50 gagaacgcag cgctctactt ctctgagatg gacaagaccg ggggcccggaa cgagcacaac
 480

55 acgggcggtg ccaagtacgg gagcggctac tgcgatgctc agtgccccgt ccagacatgg
 540

aacaacggca ccctcaacac tagccaccag ggctcgtgct gcaacgagat ggatatcctg
 600

gagggccaact cgaaggcgga agccttcacc cctcaccctt gcatcgggga caactgcgac
 660
 5 aagggcggtt gcggcttcaa ctctgtatgcg cgcggaaca aaaactactg ggcgcccgga
 720
 ggacacctcg acacctcaa gcccttcacc atggtgaccc agttcatcac ggacgacggc
 780
 10 accagctcgg gcaagcttag ccagatcgtg cgctcctacg tgcaaaacgg ccagaaggtc
 840
 gccagcgccg tgtccggcgg cgacagcatc accgtcgagg gctgctcgtc ctccgacgcc
 900
 15 tacggcgggc tcgtcgggtat gggcgaggcc ctgggcccgc gcatggtgct cgtgttcagc
 960
 atttgaacg acaacagcca gtacatgaac tggctcgaca ggggcaacgc cggcccctgc
 1020
 20 agcagcaccg agggcaaccc atccaacatc ctggccaaca accccaacac gcacgtcgtc
 1080
 25 ttctccaaca tccgctgggg agacattggg tctactacga actcgactgc gccccgccc
 1140
 ccgcctgcgt ccagcacgac gttttcgact acacggagga gctcgacgac ttcgagcagc
 1200
 30 ccgagctgca cgcagactca ctgggggcag tgcggtggca ttgggtacag cgggtgcaag
 1260
 acgtgcacgt cgggcactac gtgccagtat agcaacgact actactcgca atgcctttaa
 1320
 35
 <210> 10
 <211> 439
 40 <212> PRT
 <213> Artificial Sequence
 <220>
 <223> 154E4
 45 <400> 10
 Gln Gln Pro Gly Thr Ser Thr Pro Glu Val His Pro Lys Leu Thr Thr
 1 5 10 15
 50 Tyr Lys Cys Thr Thr Ala Gly Gly Cys Val Ala Gln Asp Thr Ser Val
 20 25 30
 55 Val Leu Asp Trp Gly Tyr His Trp Ile His Thr Val Asp Gly Tyr Thr
 35 40 45

	Ser	Cys	Thr	Thr	Ser	Ser	Gly	Val	Asp	Ser	Thr	Leu	Cys	Pro	Asp	Ala	
	50						55					60					
5	Ala	Thr	Cys	Ala	Lys	Asn	Cys	Val	Ile	Glu	Pro	Ala	Asn	Tyr	Thr	Ser	
	65					70					75					80	
10	Ala	Gly	Val	Thr	Thr	Ser	Gly	Asp	Ser	Leu	Thr	Met	Tyr	Gln	Tyr	Val	
					85					90					95		
15	Gln	Ser	Asn	Gly	Val	Tyr	Thr	Asn	Ala	Ser	Pro	Arg	Ala	Tyr	Leu	Leu	
				100					105					110			
20	Gly	Ala	Asp	Gly	Asp	Tyr	Val	Met	Phe	Lys	Leu	Leu	Asn	Gln	Glu	Leu	
			115					120					125				
25	Ser	Phe	Asp	Val	Asp	Val	Ser	Thr	Leu	Pro	Cys	Gly	Glu	Asn	Ala	Ala	
	130						135					140					
30	Leu	Tyr	Phe	Ser	Glu	Met	Asp	Lys	Thr	Gly	Gly	Arg	Asn	Glu	His	Asn	
	145					150					155					160	
35	Thr	Gly	Gly	Ala	Lys	Tyr	Gly	Ser	Gly	Tyr	Cys	Asp	Ala	Gln	Cys	Pro	
					165					170					175		
40	Val	Gln	Thr	Trp	Asn	Asn	Gly	Thr	Leu	Asn	Thr	Ser	His	Gln	Gly	Ser	
				180					185					190			
45	Cys	Cys	Asn	Glu	Met	Asp	Ile	Leu	Glu	Ala	Asn	Ser	Lys	Ala	Glu	Ala	
			195					200					205				
50	Phe	Thr	Pro	His	Pro	Cys	Ile	Gly	Asp	Asn	Cys	Asp	Lys	Gly	Gly	Cys	
	210						215					220					
55	Gly	Phe	Asn	Ser	Tyr	Ala	Arg	Gly	Asn	Lys	Asn	Tyr	Trp	Ala	Pro	Gly	
	225					230					235					240	
60	Gly	Thr	Leu	Asp	Thr	Ser	Lys	Pro	Phe	Thr	Met	Val	Thr	Gln	Phe	Ile	
					245					250					255		
65	Thr	Asp	Asp	Gly	Thr	Ser	Ser	Gly	Lys	Leu	Ser	Gln	Ile	Val	Arg	Ser	
				260					265					270			
70	Tyr	Val	Gln	Asn	Gly	Gln	Lys	Val	Ala	Ser	Ala	Val	Ser	Gly	Gly	Asp	
			275					280					285				

Ser Ile Thr Val Glu Gly Cys Ser Ser Ser Asp Ala Tyr Gly Gly Leu
 290 295 300

5

Val Gly Met Gly Glu Ala Leu Gly Arg Gly Met Val Leu Val Phe Ser
 305 310 315 320

10

Ile Trp Asn Asp Asn Ser Gln Tyr Met Asn Trp Leu Asp Ser Gly Asn
 325 330 335

15

Ala Gly Pro Cys Ser Ser Thr Glu Gly Asn Pro Ser Asn Ile Leu Ala
 340 345 350

20

Asn Asn Pro Asn Thr His Val Val Phe Ser Asn Ile Arg Trp Gly Asp
 355 360 365

25

Ile Gly Ser Thr Thr Asn Ser Thr Ala Pro Pro Pro Pro Ala Ser
 370 375 380

30

Ser Thr Thr Phe Ser Thr Thr Arg Arg Ser Ser Thr Thr Ser Ser Ser
 385 390 395 400

35

Pro Ser Cys Thr Gln Thr His Trp Gly Gln Cys Gly Gly Ile Gly Tyr
 405 410 415

40

Ser Gly Cys Lys Thr Cys Thr Ser Gly Thr Thr Cys Gln Tyr Ser Asn
 420 425 430

45

Asp Tyr Tyr Ser Gln Cys Leu
 435

<210> 11
 <211> 1320
 <212> DNA
 <213> Artificial Sequence

50

<220>
 <223> 202C12

55

<400> 11
 cagcaaccgg gtaccagcac ccccgaggtc catcccaagt tgacaacctt caagtgtaca
 60
 aagtccgggg ggtgcgtggc ccaggacacc tcggtgggtc ttgactggaa ctaccgctgg
 120
 atgcacgacg caaactacaa ctctgtgcacc gtcaacggcg gcgtcaaacac cagctctctgc
 180

cctgacgagg cgacctgtgg caagaactgc ttcacgagg gcgctgacta cgccgcctcg
 240
 5 ggcgtcacga cctcgggcag cagcctcacc ctgaggcagt tcgtcaagga cagcgagggc
 300
 aagaccaaca gcgctctctcc tcgggcctat ctctggggcg ccgacggtga ctacgtgatg
 360
 10 ttcaagctcc tcaaccagga gctgagcttc gacgtcgacg tctctacgct gccgtgtgga
 420
 gagaacgcag cgctctactt ctctgagatg gacaagaccg ggggcccggaa cgagcacaac
 480
 15 acggggcggg ccgagtacgg gagcggctac tgcgatgctc agtgccccgt ccagacatgg
 540
 aacaacggca ccctcaacac tagccaccag ggctcgtgct gcaacgagat ggatatcctg
 600
 20 gaggccaaact cgaaggcgga agccttcacc cctcaccctt gcatcgggcg caactgcgac
 660
 aagggcgggtt gcggcttcaa ctcgatgctg cgcggaaca aaaactactg ggcgccccga
 720
 ggcaccctcg acacctcaa gcccttcacc atggtgaccc agttcatcac ggacgacggc
 780
 30 accagctcgg gcaagcttag ccagatcgtg cgctcctacg tgcaaaacgg ccagaaggtc
 840
 gccagcgccg tgtccggcgg cgacagcatc accgtcgagg gctgctcgtc ctccgacgcc
 900
 35 tacggcggcc tcgtcgggtat gggcgaggcc ctgggcccgg gcatggtgct cgccatgagc
 960
 atttgaacg acgccagcgg ctcatgaac tggctcgaca gcggcgacaa cggccccctgc
 1020
 aacgagaccg agggcaaccc atccaacatc ctggccaaca accccaacac gcacgtcgtc
 1080
 45 ttctccaaca tccgctgggg agacattggg tctactacga actcgactgc gccccgccc
 1140
 ccgcctgcgt ccagcacgac gttttcgact acacggagga gctcgacgac ttcgagcagc
 1200
 50 ccgagctgca cgcagactca ctgggggcag tgcggtggca ttgggtacag cgggtgcaag
 1260
 acgtgcacgt cgggcactac gtgccagtat agcaacgact actactcgca atgccttta
 1320

<210> 12

<211> 439
 <212> PRT
 <213> Artificial Sequence

5 <220>
 <223> 202C12

<400> 12

10 Gln Gln Pro Gly Thr Ser Thr Pro Glu Val His Pro Lys Leu Thr Thr
 1 5 10 15

15 Tyr Lys Cys Thr Lys Ser Gly Gly Cys Val Ala Gln Asp Thr Ser Val
 20 25 30

20 Val Leu Asp Trp Asn Tyr Arg Trp Met His Asp Ala Asn Tyr Asn Ser
 35 40 45

Cys Thr Val Asn Gly Gly Val Asn Thr Thr Leu Cys Pro Asp Glu Ala
 50 55 60

25 Thr Cys Gly Lys Asn Cys Phe Ile Glu Gly Val Asp Tyr Ala Ala Ser
 65 70 75 80

30 Gly Val Thr Thr Ser Gly Ser Ser Leu Thr Leu Arg Gln Phe Val Lys
 85 90 95

35 Asp Ser Glu Gly Lys Thr Asn Ser Val Ser Pro Arg Ala Tyr Leu Leu
 100 105 110

40 Gly Ala Asp Gly Asp Tyr Val Met Phe Lys Leu Leu Asn Gln Glu Leu
 115 120 125

Ser Phe Asp Val Asp Val Ser Thr Leu Pro Cys Gly Glu Asn Ala Ala
 130 135 140

45 Leu Tyr Phe Ser Glu Met Asp Lys Thr Gly Gly Arg Asn Glu His Asn
 145 150 155 160

50 Thr Gly Gly Ala Glu Tyr Gly Ser Gly Tyr Cys Asp Ala Gln Cys Pro
 165 170 175

55 Val Gln Thr Trp Asn Asn Gly Thr Leu Asn Thr Ser His Gln Gly Ser
 180 185 190

Cys Cys Asn Glu Met Asp Ile Leu Glu Ala Asn Ser Lys Ala Glu Ala
 195 200 205

5 Phe Thr Pro His Pro Cys Ile Gly Asp Asn Cys Asp Lys Gly Gly Cys
 210 215 220
 Gly Phe Asn Ser Tyr Ala Arg Gly Asn Lys Asn Tyr Trp Ala Pro Gly
 225 230 235 240
 10 Gly Thr Leu Asp Thr Ser Lys Pro Phe Thr Met Val Thr Gln Phe Ile
 245 250 255
 15 Thr Asp Asp Gly Thr Ser Ser Gly Lys Leu Ser Gln Ile Val Arg Ser
 260 265 270
 20 Tyr Val Gln Asn Gly Gln Lys Val Ala Ser Ala Val Ser Gly Gly Asp
 275 280 285
 25 Ser Ile Thr Val Glu Gly Cys Ser Ser Ser Asp Ala Tyr Gly Gly Leu
 290 295 300
 Val Gly Met Gly Glu Ala Leu Gly Arg Gly Met Val Leu Ala Met Ser
 305 310 315 320
 30 Ile Trp Asn Asp Ala Ser Gly Phe Met Asn Trp Leu Asp Ser Gly Asp
 325 330 335
 35 Asn Gly Pro Cys Asn Glu Thr Glu Gly Asn Pro Ser Asn Ile Leu Ala
 340 345 350
 40 Asn Asn Pro Asn Thr His Val Val Phe Ser Asn Ile Arg Trp Gly Asp
 355 360 365
 45 Ile Gly Ser Thr Thr Asn Ser Thr Ala Pro Pro Pro Pro Ala Ser
 370 375 380
 Ser Thr Thr Phe Ser Thr Thr Arg Arg Ser Ser Thr Thr Ser Ser Ser
 385 390 395 400
 50 Pro Ser Cys Thr Gln Thr His Trp Gly Gln Cys Gly Gly Ile Gly Tyr
 405 410 415
 55 Ser Gly Cys Lys Thr Cys Thr Ser Gly Thr Thr Cys Gln Tyr Ser Asn
 420 425 430
 Asp Tyr Tyr Ser Gln Cys Leu

435

5 <210> 13
 <211> 1323
 <212> DNA
 <213> Artificial Sequence

10 <220>
 <223> 272A9

15 <400> 13
 cagcaaccgg gtaccagcac ccccgaggtc catcccaagt tgacaaccta caagtgtaca
 60

20 aagtccgggg ggtgcggtggc ccaggacacc tcggtggtcc ttgactggaa ctaccgctgg
 120

25 atgcacgacg caaactacaa ctcgatgcacc gtcaacggcg gcgtcaacac cacgctctgc
 180

30 cctgacgcgg cgacctgtgc gaagaactgc gtgatcgagc cggccaacta caccagcgcc
 240

35 ggcgtcacga cctcggggcga cagcctcacc atgtaccagt acgttcagag caacggcgctc
 300

40 tacaccaacg cctctcctcg gctgtatctc ctggggccccg acaagaacta cgtgatgctg
 360

45 aagctcctag gccaggagct gagcttcgac gtcgacctct ctgctctgcc gtgtggagag
 420

50 aacggcgccc tctacctgtc tgaaatgagc gccaccgggg gccgcaacga atataacacg
 480

55 ggcggtgccg agtacgggag cggctactgc gatgctcagt gcccgcgtcat cgccctggaag
 540

60 aacggcaccc tcaacactag cggcgcaagc tactgctgca acgagatgga tatcctggag
 600

65 gccaaactcga gggcgaattc gtacaccctt cacccttgca gtgccacgga ctgcgacaag
 660

70 ggcggttgcg gcttcaaccc ctatgctctc ggccaaaaaa gctactgggg ccccgagggc
 720

75 accgttgaca cctccaagcc cttcaccatc accaccaggt tcatcacgaa cgacggcacc
 780

80 accaccggca ccctttccga aatccggcgc cagtacatgc aaaacggcaa ggtgatcgcc
 840

85 aatgccgttt cctccactgg cgtcaactcc atcaccgagg actggtgcac gtccgtcgac
 900

90 ggctcggccg ccacctttgg cggcctcacc accatgggca aggccctggg ccgcggcatg
 960

gtgctcatct tcagcatttg gaacgacgcc agcgggttta tgaactggct cgacagcggc
 1020

5 aacgccggcc cctgcagcag caccgagggc aaccagact tgatcaaggc ccagaacccc
 1080

acgacgcacg tcgtcttctc caacatccgc tggggagaca ttgggtctac tttcaagggt
 1140

10 tctgatggct cggtgacgac gacgacgtcg actacatcga ccaagaccac gacttcgacc
 1200

15 gcgccggggc caacgcagac tcactatggg cagtgcggtg gccaaagggtg gactggggccc
 1260

acggcttgcg catcgcccta cacgtgccag gttctgaacc cgtgggtactc gcaatgcctt
 1320

20 taa
 1323

25 <210> 14
 <211> 440
 <212> PRT
 <213> Artificial Sequence

30 <220>
 <223> 272A9
 <400> 14

35 Gln Gln Pro Gly Thr Ser Thr Pro Glu Val His Pro Lys Leu Thr Thr
 1 5 10 15

Tyr Lys Cys Thr Lys Ser Gly Gly Cys Val Ala Gln Asp Thr Ser Val
 20 25 30

40 Val Leu Asp Trp Asn Tyr Arg Trp Met His Asp Ala Asn Tyr Asn Ser
 35 40 45

45 Cys Thr Val Asn Gly Gly Val Asn Thr Thr Leu Cys Pro Asp Ala Ala
 50 55 60

50 Thr Cys Ala Lys Asn Cys Val Ile Glu Pro Ala Asn Tyr Thr Ser Ala
 65 70 75 80

55 Gly Val Thr Thr Ser Gly Asp Ser Leu Thr Met Tyr Gln Tyr Val Gln
 85 90 95

Ser Asn Gly Val Tyr Thr Asn Ala Ser Pro Arg Leu Tyr Leu Leu Gly
 100 105 110

5

10

15

20

25

30

35

40

45

50

55

	340	345	350
5	Asp Leu Ile Lys Ala Gln Asn Pro Thr Thr His Val Val Phe Ser Asn 355 360 365		
10	Ile Arg Trp Gly Asp Ile Gly Ser Thr Phe Lys Gly Ser Asp Gly Ser 370 375 380		
15	Val Thr Thr Thr Thr Ser Thr Thr Ser Thr Lys Thr Thr Thr Ser Thr 385 390 395 400		
20	Ala Pro Gly Pro Thr Gln Thr His Tyr Gly Gln Cys Gly Gly Gln Gly 405 410 415		
25	Trp Thr Gly Pro Thr Ala Cys Ala Ser Pro Tyr Thr Cys Gln Val Leu 420 425 430		
30	Asn Pro Trp Tyr Ser Gln Cys Leu 435 440		
35	<210> 15 <211> 1317 <212> DNA <213> Artificial Sequence		
40	<220> <223> 278F10		
45	<400> 15 cagcaaccgg gtaccagcac ccccgaggtc catcccaagt tgacaacctt caagtgtaca 60		
50	aagtccgggg ggtgcgtgga gcaggacacc tcggtgggtcc ttgactggaa ctaccgctgg 120		
55	ttccacacct cggacaacac cacctcgtgc accacgtcct cgggcatoga ctgcagctgc 180		
60	tgccctgacg cagcgacctg tgccgagaac tgcgtcgtgg agggcaccga ctacaccagc 240		
65	tcgggcatcg agacctcggg cagcagcctc accctgaggc agttcgtcaa ggacagcgag 300		
70	ggcaagacca acagcgtctc tcctcggggc tatctcctgg gcgcgcagcg tgactacgtg 360		
75	atgttcaagc tcctcaacca ggagctgagc ttcgacgtcg acgtctctac gctgccgtgt 420		
80	ggagagaacg cagcgtctta cttctctgag atggacaaga ccggggggccg gaacgagcac 480		

aacacgggcg gtgccaagta cgggagcggc tactgcatg ctcaagtccc cgtccagaca
 540

5 tggaacaacg gcaccctcaa cactagccac cagggctcgt gctgcaacga gatggatatac
 600

ctggaggcca actcgaaggc ggaagccttc acccctcacc cctgcatcgg cgacaactgc
 660

10 gacaagggcg gttgcggtt caactcgtat gcgcgcggca aaaaaacta ctgggcgccc
 720

ggaggcacc tcgacacctc caagcccttc accatggtga cccagttcat cacggacgac
 780

15 ggcaccagct cgggcaacct tgtgagcatc acccgcaagt accagcaaaa cggcgtcgac
 840

20 atccccagcg cccagcccg cggcgacacc atctcgtcct gcccgtcgc ctcagcctac
 900

ggcggcctcg ccaccatggg caaggccctg agcagcggca tgggtgctcgt gttcagcatt
 960

25 tggaacgaca acagccagta catgaactgg ctcgacagcg gcaacgcgg cccctgcagc
 1020

agcaccgagg gcaaccatc caacatcctg gccacaacc ccaacacgca cgtcgtcttc
 1080

30 tccaacatcc gctggggaga cattgggtct actacgaact cgactgcgcc cccgccccg
 1140

35 cctgcgtcca gcacgacgtt ttcgactaca cggaggagct cgacgacttc gagcagcccc
 1200

agctgcacgc agactcactg ggggcagtgc ggtggcattg ggtacagcgg gtgcaagacg
 1260

40 tgcacgtcgg gcactacgtg ccagtatagc aacgactact actcgcaatg cctttaa
 1317

45 <210> 16
 <211> 438
 <212> PRT
 <213> Artificial Sequence

50 <220>
 <223> 278F10
 <400> 16

55 Gln Gln Pro Gly Thr Ser Thr Pro Glu Val His Pro Lys Leu Thr Thr
 1 5 10 15

Tyr Lys Cys Thr Lys Ser Gly Gly Cys Val Glu Gln Asp Thr Ser Val

	20							25					30				
5	Val	Leu	Asp	Trp	Asn	Tyr	Arg	Trp	Phe	His	Thr	Ser	Asp	Asn	Thr	Thr	
			35					40					45				
10	Ser	Cys	Thr	Thr	Ser	Ser	Gly	Ile	Asp	Ser	Thr	Leu	Cys	Pro	Asp	Ala	
		50					55					60					
15	Ala	Thr	Cys	Ala	Glu	Asn	Cys	Val	Val	Glu	Gly	Thr	Asp	Tyr	Thr	Ser	
	65					70					75					80	
20	Ser	Gly	Ile	Glu	Thr	Ser	Gly	Ser	Ser	Leu	Thr	Leu	Arg	Gln	Phe	Val	
					85					90					95		
25	Lys	Asp	Ser	Glu	Gly	Lys	Thr	Asn	Ser	Val	Ser	Pro	Arg	Ala	Tyr	Leu	
				100					105					110			
30	Leu	Gly	Ala	Asp	Gly	Asp	Tyr	Val	Met	Phe	Lys	Leu	Leu	Asn	Gln	Glu	
			115					120					125				
35	Leu	Ser	Phe	Asp	Val	Asp	Val	Ser	Thr	Leu	Pro	Cys	Gly	Glu	Asn	Ala	
		130					135					140					
40	Ala	Leu	Tyr	Phe	Ser	Glu	Met	Asp	Lys	Thr	Gly	Gly	Arg	Asn	Glu	His	
	145					150					155					160	
45	Asn	Thr	Gly	Gly	Ala	Lys	Tyr	Gly	Ser	Gly	Tyr	Cys	Asp	Ala	Gln	Cys	
					165					170					175		
50	Pro	Val	Gln	Thr	Trp	Asn	Asn	Gly	Thr	Leu	Asn	Thr	Ser	His	Gln	Gly	
				180					185					190			
55	Ser	Cys	Cys	Asn	Glu	Met	Asp	Ile	Leu	Glu	Ala	Asn	Ser	Lys	Ala	Glu	
			195					200					205				
60	Ala	Phe	Thr	Pro	His	Pro	Cys	Ile	Gly	Asp	Asn	Cys	Asp	Lys	Gly	Gly	
		210					215					220					
65	Cys	Gly	Phe	Asn	Ser	Tyr	Ala	Arg	Gly	Asn	Lys	Asn	Tyr	Trp	Ala	Pro	
	225					230					235					240	
70	Gly	Gly	Thr	Leu	Asp	Thr	Ser	Lys	Pro	Phe	Thr	Met	Val	Thr	Gln	Phe	
					245					250					255		

[illegible]

aagtccgggg ggtgcggtgga gcaggacacc tcggtggtcc ttgactggaa ctaccgctgg
 120

5 ttccacacct cggacaacac cacctcgtgc accacgtcct cgggcatoga ctcgacgctc
 180

tgccctgacg cagcgacctg tgccgagaac tgcgtcgtgg agggcacoga ctacaccagc
 240

10 tcgggcatcg agacctcggg cagcagcctc accctgaggg agttcgtcaa ggacagcgag
 300

15 ggcaagacca acagcgtctc tcctcggggc tatctcctgg gcgccgacgg tgactacgtg
 360

atgttcaagc tcctcaacca ggagctgagc ttcgacgtcg acgtctctac gctgccgtgt
 420

20 ggagagaacg cagcgctcta cttctctgag atggacaaga cggggggccg gaacgagcac
 480

aacacgggcg gtgccaagta cgggagcggc tactgcatg ctcagtgcc cgtccagaca
 540

25 tggaacaacg gcaccctcaa cactagccac cagggtcgt gctgcaacga gatggatata
 600

30 ctggaggcca actcgaaggc ggaagccttc acccctcacc cctgcatcgg cgacaactgc
 660

gacaagggcg gttgcggtt caactcgtat gcgcgcgga acaaaaacta ctgggcgccc
 720

35 ggaggcacc tcgacacctc caagcccttc accatggtga ccagttcat cagggacgac
 780

ggaccagct cgggcaagct tagccagatc gtgcgtcct acgtgcaaaa cggccagaag
 840

40 gtcgccagcg ccgtgtccgg cggcgacagc atcaccgtcg agggctgctc gtctccgac
 900

45 gcctacggcg gcctcgtcgg tatgggagag gccctgggccc gggcatggt gctcgtgttc
 960

agcatttga acgacaacag ccagtacatg aactggctcg acagcggcaa cgccggcccc
 1020

50 tgcagcagca ccgagggcaa cccatccaac atcctggcca acaaccccaa cagcgacgtc
 1080

gtcttctcca acatccgctg gggagacatt gggcttacta cgaactcgac tgcgccccg
 1140

55 ccccgctg cgtccagcac gacgttttcg actacacgga ggagctcgac gacttcgagc
 1200

agcccgagct gcacgcagac tcactggggg cagtgcggtg gcattgggta cagcgggtgc
1260

5 aagacgtgca cgtcgggcac tacgtgccag tatagcaacg actactactc gcaatgcctt
1320

taa
1323

10 <210> 18
<211> 440
<212> PRT
<213> Artificial Sequence

15 <220>
<223> 293B2

<400> 18

20 Gln Gln Pro Gly Ser Val Thr Pro Glu Val His Pro Lys Leu Pro Thr
1 5 10 15

25 Trp Gln Cys Thr Lys Ser Gly Gly Cys Val Glu Gln Asp Thr Ser Val
20 25 30

30 Val Leu Asp Trp Asn Tyr Arg Trp Phe His Thr Ser Asp Asn Thr Thr
35 40 45

35 Ser Cys Thr Thr Ser Ser Gly Ile Asp Ser Thr Leu Cys Pro Asp Ala
50 55 60

Ala Thr Cys Ala Glu Asn Cys Val Val Glu Gly Thr Asp Tyr Thr Ser
65 70 75 80

40 Ser Gly Ile Glu Thr Ser Gly Ser Ser Leu Thr Leu Arg Gln Phe Val
85 90 95

45 Lys Asp Ser Glu Gly Lys Thr Asn Ser Val Ser Pro Arg Ala Tyr Leu
100 105 110

50 Leu Gly Ala Asp Gly Asp Tyr Val Met Phe Lys Leu Leu Asn Gln Glu
115 120 125

55 Leu Ser Phe Asp Val Asp Val Ser Thr Leu Pro Cys Gly Glu Asn Ala
130 135 140

Ala Leu Tyr Phe Ser Glu Met Asp Lys Thr Gly Gly Arg Asn Glu His
145 150 155 160

	Asn	Thr	Gly	Gly	Ala	Lys	Tyr	Gly	Ser	Gly	Tyr	Cys	Asp	Ala	Gln	Cys	
					165					170					175		
5	Pro	Val	Gln	Thr	Trp	Asn	Asn	Gly	Thr	Leu	Asn	Thr	Ser	His	Gln	Gly	
				180					185					190			
10	Ser	Cys	Cys	Asn	Glu	Met	Asp	Ile	Leu	Glu	Ala	Asn	Ser	Lys	Ala	Glu	
			195					200					205				
15	Ala	Phe	Thr	Pro	His	Pro	Cys	Ile	Gly	Asp	Asn	Cys	Asp	Lys	Gly	Gly	
		210					215					220					
20	Cys	Gly	Phe	Asn	Ser	Tyr	Ala	Arg	Gly	Asn	Lys	Asn	Tyr	Trp	Ala	Pro	
	225					230					235					240	
25	Gly	Gly	Thr	Leu	Asp	Thr	Ser	Lys	Pro	Phe	Thr	Met	Val	Thr	Gln	Phe	
					245					250					255		
30	Ile	Thr	Asp	Asp	Gly	Thr	Ser	Ser	Gly	Lys	Leu	Ser	Gln	Ile	Val	Arg	
				260					265					270			
35	Ser	Tyr	Val	Gln	Asn	Gly	Gln	Lys	Val	Ala	Ser	Ala	Val	Ser	Gly	Gly	
			275					280					285				
40	Asp	Ser	Ile	Thr	Val	Glu	Gly	Cys	Ser	Ser	Ser	Asp	Ala	Tyr	Gly	Gly	
		290					295					300					
45	Leu	Val	Gly	Met	Gly	Glu	Ala	Leu	Gly	Arg	Gly	Met	Val	Leu	Val	Phe	
	305					310					315					320	
50	Ser	Ile	Trp	Asn	Asp	Asn	Ser	Gln	Tyr	Met	Asn	Trp	Leu	Asp	Ser	Gly	
				325						330					335		
55	Asn	Ala	Gly	Pro	Cys	Ser	Ser	Thr	Glu	Gly	Asn	Pro	Ser	Asn	Ile	Leu	
				340					345					350			
60	Ala	Asn	Asn	Pro	Asn	Thr	His	Val	Val	Phe	Ser	Asn	Ile	Arg	Trp	Gly	
			355					360					365				
65	Asp	Ile	Gly	Ser	Thr	Thr	Asn	Ser	Thr	Ala	Pro	Pro	Pro	Pro	Pro	Ala	
		370					375					380					
70	Ser	Ser	Thr	Thr	Phe	Ser	Thr	Thr	Arg	Arg	Ser	Ser	Thr	Thr	Ser	Ser	
	385					390					395					400	

5 Ser Pro Ser Cys Thr Gln Thr His Trp Gly Gln Cys Gly Gly Ile Gly
 405 410 415

Tyr Ser Gly Cys Lys Thr Cys Thr Ser Gly Thr Thr Cys Gln Tyr Ser
 420 425 430

10 Asn Asp Tyr Tyr Ser Gln Cys Leu
 435 440

15 <210> 19
 <211> 1323
 <212> DNA
 <213> Artificial Sequence

20 <220>
 <223> 309A11

25 <400> 19
 cagcaaccgg gttcgggtgac ccccgaggtc catcccaagt tgcccacctg gcagtgtaca
 60

aagtccgggg ggtgcggtgga gcaggacacc tcggtggtcc ttgactggaa ctaccgctgg
 120

30 ttccacacct cggacaacac cacctcgtgc accacgtcct cgggcatcga ctgcacgctc
 180

tgccctgacg cagcgacctg tgccgagaac tgcgtcgtgg agggcaccga ctacaccagc
 240

35 tcgggcatcg agacctcggg cagcagcctc accctgaggc agttcgtcaa ggacagcgag
 300

40 ggcaagacca acagcgtctc tcctcggggc tatctcctgg gcgccgacgg tgactacgtg
 360

atgttcaagc tcctcaacca ggagctgagc ttcgacgtcg acgtctctac gctgccgtgt
 420

45 ggagagaacg cagcgctcta cttctctgag atggacaaga ccggggggccg gaacgagcac
 480

aacacggggc gtgccaagta cgggagcggc tactgcgatg ctcagtgcc cgtccagaca
 540

50 tggaacaacg gcaccctcaa cactagccac cagggtcgt gctgcaacga gatggatatc
 600

55 ctggaggcca actcgaaggc ggaagccttc acccctcacc cctgcacgg cgacaactgc
 660

gacaagggcg gttgcggctt caactcgtat gcgcgcggca acaaaaacta ctgggcgccc
 720

ggaggcacc tcgacacctc caagcccttc accatggtga cccagttcat caccgacgac
 780
 5 ggcaccagct cgggcaagct tagccagatc gtgcgctcct acgtgcaaaa cggccagaag
 840
 gtcgccagcg ccgtgtccgg cggcgacagc atcacgctcg agggctgctc gtccgcctca
 900
 10 gcctacggcg gcctcgccac catgggcaag gccctgagca ggggcatggt gctcgtgttc
 960
 agcatttga acgacaacag ccagtacatg aactggctcg acagcggcaa cgccggcccc
 1020
 15 tgcagcagca ccgagggcaa cccatccaac atcctggcca acaaccccaa caccgacgtc
 1080
 gtcttctcca acatccgctg gggagacatt gggcttacta cgaactcgac tgcgcccccg
 1140
 20 ccccgccctg cgtccagcac gacgttttcg actacacgga ggagctcgac gacttcgagc
 1200
 25 agcccgagct gcacgcagac tcaactggggg cagtgcggtg gcattgggta cagcgggtgc
 1260
 aagacgtgca cgtcgggcac tacgtgccag tatagcaacg actactactc gcaatgcctt
 1320
 30 taa
 1323
 35 <210> 20
 <211> 440
 <212> PRT
 <213> Artificial Sequence
 40 <220>
 <223> 309A1
 <400> 20
 45 Gln Gln Pro Gly Ser Val Thr Pro Glu Val His Pro Lys Leu Pro Thr
 1 5 10 15
 50 Trp Gln Cys Thr Lys Ser Gly Gly Cys Val Glu Gln Asp Thr Ser Val
 20 25 30
 Val Leu Asp Trp Asn Tyr Arg Trp Phe His Thr Ser Asp Asn Thr Thr
 35 40 45
 55 Ser Cys Thr Thr Ser Ser Gly Ile Asp Ser Thr Leu Cys Pro Asp Ala
 50 55 60

	Ala	Thr	Cys	Ala	Glu	Asn	Cys	Val	Val	Glu	Gly	Thr	Asp	Tyr	Thr	Ser	
	65					70					75					80	
5																	
	Ser	Gly	Ile	Glu	Thr	Ser	Gly	Ser	Ser	Leu	Thr	Leu	Arg	Gln	Phe	Val	
				85						90					95		
10																	
	Lys	Asp	Ser	Glu	Gly	Lys	Thr	Asn	Ser	Val	Ser	Pro	Arg	Ala	Tyr	Leu	
				100					105					110			
15																	
	Leu	Gly	Ala	Asp	Gly	Asp	Tyr	Val	Met	Phe	Lys	Leu	Leu	Asn	Gln	Glu	
			115					120					125				
20																	
	Leu	Ser	Phe	Asp	Val	Asp	Val	Ser	Thr	Leu	Pro	Cys	Gly	Glu	Asn	Ala	
		130					135					140					
25																	
	Ala	Leu	Tyr	Phe	Ser	Glu	Met	Asp	Lys	Thr	Gly	Gly	Arg	Asn	Glu	His	
	145					150					155					160	
30																	
	Asn	Thr	Gly	Gly	Ala	Lys	Tyr	Gly	Ser	Gly	Tyr	Cys	Asp	Ala	Gln	Cys	
					165					170					175		
35																	
	Pro	Val	Gln	Thr	Trp	Asn	Asn	Gly	Thr	Leu	Asn	Thr	Ser	His	Gln	Gly	
				180					185					190			
40																	
	Ser	Cys	Cys	Asn	Glu	Met	Asp	Ile	Leu	Glu	Ala	Asn	Ser	Lys	Ala	Glu	
			195					200					205				
45																	
	Ala	Phe	Thr	Pro	His	Pro	Cys	Ile	Gly	Asp	Asn	Cys	Asp	Lys	Gly	Gly	
		210					215					220					
50																	
	Cys	Gly	Phe	Asn	Ser	Tyr	Ala	Arg	Gly	Asn	Lys	Asn	Tyr	Trp	Ala	Pro	
	225					230					235					240	
55																	
	Gly	Gly	Thr	Leu	Asp	Thr	Ser	Lys	Pro	Phe	Thr	Met	Val	Thr	Gln	Phe	
					245					250					255		
60																	
	Ile	Thr	Asp	Asp	Gly	Thr	Ser	Ser	Gly	Lys	Leu	Ser	Gln	Ile	Val	Arg	
				260					265					270			
65																	
	Ser	Tyr	Val	Gln	Asn	Gly	Gln	Lys	Val	Ala	Ser	Ala	Val	Ser	Gly	Gly	
			275					280					285				
70																	
	Asp	Ser	Ile	Thr	Val	Glu	Gly	Cys	Ser	Ser	Ala	Ser	Ala	Tyr	Gly	Gly	
		290					295					300					

5 Leu Ala Thr Met Gly Lys Ala Leu Ser Ser Gly Met Val Leu Val Phe
 305 310 315 320

10 Ser Ile Trp Asn Asp Asn Ser Gln Tyr Met Asn Trp Leu Asp Ser Gly
 325 330 335

15 Asn Ala Gly Pro Cys Ser Ser Thr Glu Gly Asn Pro Ser Asn Ile Leu
 340 345 350

20 Ala Asn Asn Pro Asn Thr His Val Val Phe Ser Asn Ile Arg Trp Gly
 355 360 365

25 Asp Ile Gly Ser Thr Thr Asn Ser Thr Ala Pro Pro Pro Pro Pro Ala
 370 375 380

30 Ser Ser Thr Thr Phe Ser Thr Thr Arg Arg Ser Ser Thr Thr Ser Ser
 385 390 395 400

35 Ser Pro Ser Cys Thr Gln Thr His Trp Gly Gln Cys Gly Gly Ile Gly
 405 410 415

40 Tyr Ser Gly Cys Lys Thr Cys Thr Ser Gly Thr Thr Cys Gln Tyr Ser
 420 425 430

45 Asn Asp Tyr Tyr Ser Gln Cys Leu
 435 440

50 <210> 21
 <211> 1320
 <212> DNA
 <213> Artificial Sequence

55 <220>
 <223> 11G8

60 <400> 21
 cagcaaccgg gtaccagcac ccccgaggtc catcccaagt tgacaaccta caagtgtaca

65 aagtccgggg ggtgctgtgg ccaggacacc tcggtggtcc ttgactggaa ctaccgctgg

70 atgcacgacg caaactacaa ctctgtgcacc gtcaacggcg gcgtcaacac cacgctctgc

75 cctgacgagg cgacctgtgg caagaactgc ttcacgagg gcgtcgacta cgcgcgctcg

80 240

ggcgtcacga cctcgggcag cagcctcacc ctgaggcagt tcgtcaagga cagcgagggc
 300
 5 aagaccaaca gcgtctctcc tcgggcctat ctctggggcg ccgacgggtga ctacgtgatg
 360
 ttcaagctcc tcaaccagga gctgagcttc gacgtcgacg tctctacgct gccgtgtgga
 420
 10 gagaacgcag cgctctactt ctctgagatg gacaagaccg ggggccgcaa cgaatataac
 480
 acgggcggtg ccgagtacgg gagcggctac tgcgatgctc agtgccccgt ccagacatgg
 540
 15 aacaacggca ccctcaacac tagccaccag ggctcgtgct gcaacgagat ggatatcctg
 600
 20 gaggccaact cgaaggcgga agccttcacc cctcaccctt gcatcgggca caactgcgac
 660
 aagggcggtt gcggcttcaa ctcgatgctg cgcggcaaca aaaactactg ggcgcccgga
 720
 25 ggcaccctcg acacctccaa gcccttcacc atggtgaccc agttcatcac ggacgacggc
 780
 accagctcgg gcaagcttag ccagatcgtg cgctcctacg tgcaaaacgg ccagaaggtc
 840
 30 gccagcgccg tgtccggcgg cgacagcatc accgtcgagg gctgctcgtc ctccgacgcc
 900
 tacggcggcc tcgtcgggtat gggcgaggcc ctgggccgcg gcatggtgct cgtgttcagc
 960
 35 atttgaacg acaacagcca gtacatgaac tggctcgaca gcggcaacgc cggccccctgc
 1020
 40 agcagcaccg agggcaaccc atccaacatc ctggccaaca accccaacac gcacgtcgtc
 1080
 ttctccaaca tccgctgggg agacattggg tctactacga actcgactgc gcccccgccc
 1140
 45 ccgcctgcgt ccagcacgac gttttcgact acacggagga gctcgacgac ttcgagcagc
 1200
 50 ccgagctgca cgcagactca ctgggggcag tgcggtggca ttgggtacag cgggtgcaag
 1260
 acgtgcacgt cgggcactac gtgccagtat agcaacgact actactcgca atgcctttaa
 1320
 55
 <210> 22
 <211> 439
 <212> PRT
 <213> Artificial Sequence

<220>

<223> 11G8

5 <400> 22

Gln Gln Pro Gly Thr Ser Thr Pro Glu Val His Pro Lys Leu Thr Thr
 1 5 10 15

10 Tyr Lys Cys Thr Lys Ser Gly Gly Cys Val Ala Gln Asp Thr Ser Val
 20 25 30

15 Val Leu Asp Trp Asn Tyr Arg Trp Met His Asp Ala Asn Tyr Asn Ser
 35 40 45

20 Cys Thr Val Asn Gly Gly Val Asn Thr Thr Leu Cys Pro Asp Glu Ala
 50 55 60

25 Thr Cys Gly Lys Asn Cys Phe Ile Glu Gly Val Asp Tyr Ala Ala Ser
 65 70 75 80

Gly Val Thr Thr Ser Gly Ser Ser Leu Thr Leu Arg Gln Phe Val Lys
 85 90 95

30 Asp Ser Glu Gly Lys Thr Asn Ser Val Ser Pro Arg Ala Tyr Leu Leu
 100 105 110

35 Gly Ala Asp Gly Asp Tyr Val Met Phe Lys Leu Leu Asn Gln Glu Leu
 115 120 125

40 Ser Phe Asp Val Asp Val Ser Thr Leu Pro Cys Gly Glu Asn Ala Ala
 130 135 140

45 Leu Tyr Phe Ser Glu Met Asp Lys Thr Gly Gly Arg Asn Glu Tyr Asn
 145 150 155 160

Thr Gly Gly Ala Glu Tyr Gly Ser Gly Tyr Cys Asp Ala Gln Cys Pro
 165 170 175

50 Val Gln Thr Trp Asn Asn Gly Thr Leu Asn Thr Ser His Gln Gly Ser
 180 185 190

55 Cys Cys Asn Glu Met Asp Ile Leu Glu Ala Asn Ser Lys Ala Glu Ala
 195 200 205

Phe Thr Pro His Pro Cys Ile Gly Asp Asn Cys Asp Lys Gly Gly Cys

<210> 23
 <211> 1317
 <212> DNA
 <213> Artificial Sequence

5

<220>
 <223> 92A12

10

<400> 23
 cagcaaccgg gtaccagcac ccccgaggtc catcccaagt tgacaaccta caagtgtaca
 60

aagtccgggg ggtgcgtggc ccaggacacc tcggtgggtcc ttgactggaa ctaccgctgg
 120

15

atgcacgacg caaactacaa ctctgtgcacc gtcaacggcg gcgtcaacac cacgctctgc
 180

20

cctgacgcgg cgacctgtgc gaagaactgc gtgatcgagc cggccaacta caccagcgcc
 240

ggcgtcacga cctcggggcg cagcctcacc atgtaccagt acgttcagag caacggcgctc
 300

25

tacaccaacg cctctcctcg ggcctatctc ctgggcgccg acggtgacta cgtgatgttc
 360

aagctcctca accaggagct gagcttcgac gtcgacgtct ctacgctgcc gtgtggagag
 420

30

aacgcagcgc tctacttctc tgagatggac aagaccgggg gccggaacga gcacaacacg
 480

35

ggcggtgcc agtacgggag cggctactgc gatgctcagt gccccgtcca gacatggaac
 540

aacggcaccc tcaacactag ccaccagggc tcgtgctgca acgagatgga tatcctggag
 600

40

gccaaactcga aggcggaagc cttcaccct caccctgca tcggcgacaa ctgcgacaag
 660

ggcggttgcg gcttcaactc gtatgcgcgc ggcaacaaaa actactgggc gcccgagggc
 720

45

accctcgaca cctccaagcc cttcaccatg gtgaccagcgc tcacacgga cgacggcacc
 780

50

agctcgggca agcttagcca gatcgtgcgc tcctacgtgc aaaacggcca gaaggctgcc
 840

agcgccgtgt ccggcgggcg cagcatcacc gtcgagggct gctcgtctct cgacgcctac
 900

55

ggcggcctcg tcggtatggg cgaggccctg ggccgcgga tgggtgctcgt gttcagcatt
 960

tggaacgaca acagccagta catgaactgg ctcgacagcg gcaacgcccg cccctgcagc
 1020

55 Ala Asp Gly Asp Tyr Val Met Phe Lys Leu Leu Asn Gln Glu Leu Ser
115 120 125

	Phe	Asp	Val	Asp	Val	Ser	Thr	Leu	Pro	Cys	Gly	Glu	Asn	Ala	Ala	Leu	
	130						135					140					
5	Tyr	Phe	Ser	Glu	Met	Asp	Lys	Thr	Gly	Gly	Arg	Asn	Glu	His	Asn	Thr	
	145					150					155					160	
10	Gly	Gly	Ala	Lys	Tyr	Gly	Ser	Gly	Tyr	Cys	Asp	Ala	Gln	Cys	Pro	Val	
					165					170					175		
15	Gln	Thr	Trp	Asn	Asn	Gly	Thr	Leu	Asn	Thr	Ser	His	Gln	Gly	Ser	Cys	
				180					185					190			
20	Cys	Asn	Glu	Met	Asp	Ile	Leu	Glu	Ala	Asn	Ser	Lys	Ala	Glu	Ala	Phe	
			195					200					205				
25	Thr	Pro	His	Pro	Cys	Ile	Gly	Asp	Asn	Cys	Asp	Lys	Gly	Gly	Cys	Gly	
		210					215					220					
30	Thr	Leu	Asp	Thr	Ser	Lys	Pro	Phe	Thr	Met	Val	Thr	Gln	Phe	Ile	Thr	
					245					250					255		
35	Asp	Asp	Gly	Thr	Ser	Ser	Gly	Lys	Leu	Ser	Gln	Ile	Val	Arg	Ser	Tyr	
				260					265					270			
40	Val	Gln	Asn	Gly	Gln	Lys	Val	Ala	Ser	Ala	Val	Ser	Gly	Gly	Asp	Ser	
			275					280					285				
45	Ile	Thr	Val	Glu	Gly	Cys	Ser	Ser	Ser	Asp	Ala	Tyr	Gly	Gly	Leu	Val	
		290					295					300					
50	Gly	Met	Gly	Glu	Ala	Leu	Gly	Arg	Gly	Met	Val	Leu	Val	Phe	Ser	Ile	
	305					310					315					320	
55	Trp	Asn	Asp	Asn	Ser	Gln	Tyr	Met	Asn	Trp	Leu	Asp	Ser	Gly	Asn	Ala	
					325					330					335		
60	Gly	Pro	Cys	Ser	Ser	Thr	Glu	Gly	Asn	Pro	Ser	Asn	Ile	Leu	Ala	Asn	
				340					345					350			
65	Asn	Pro	Asn	Thr	His	Val	Val	Phe	Ser	Asn	Ile	Arg	Trp	Gly	Asp	Ile	
			355					360					365				

Gly Ser Thr Thr Asn Ser Thr Ala Pro Pro Pro Pro Pro Ala Ser Ser
 370 375 380

5

Thr Thr Phe Ser Thr Thr Arg Arg Ser Ser Thr Thr Ser Ser Ser Pro
 385 390 395 400

10

Ser Cys Thr Gln Thr His Trp Gly Gln Cys Gly Gly Ile Gly Tyr Ser
 405 410 415

15

Gly Cys Lys Thr Cys Thr Ser Gly Thr Thr Cys Gln Tyr Ser Asn Asp
 420 425 430

20

Tyr Tyr Ser Gln Cys Leu
 435

25

<210> 25
 <211> 1320
 <212> DNA
 <213> Artificial Sequence

30

<220>
 <223> 240H12

35

<400> 25
 cagcaaccgg gtaccagcac ccccgaggtc catcccaagt tgacaaccta caagtgtaca
 60

actgctgggg ggtgctggc ccaggacacc tcggtggtcc ttgactgggg ctaccactgg
 120

attcacacgg tcgatgggta cacatcgtgc accacatcgt cgggcgtcga cagcacgctc
 180

40

tgccctgacg cggcgacctg tgccaagaac tgcgtgatcg agccggccaa ctacaccagc
 240

45

gccggcgtca cgacctcggg cgacagcctc accatgtacc agtacgttca gagcaacggc
 300

gtctacacca acgcctctcc tcgggcctat ctctggggcg ccgacgggtga ctacgtgatg
 360

50

ttcaagctcc tcaaccagga gctgagcttc gacgtcgacg tctctacgct gccgtgtgga
 420

gagaacgcag cgctctactt ctctgagatg gacaagaccg ggggcgggaa cgagcacaac
 480

55

acgggcggtg ccaagtacgg gagcggctac tgcgatgctc agtgccccgt ccagacatgg
 540

aacaacggca ccctcaacac tagccaccag ggctcgtgct gcaacgagat ggatatcctg
 600

gagggccaact cgaaggcgga agccttcacc cctcaccctt gcatcgggga caactgcgac
 660

5 aagggcggtt gcggcttcaa ctcgtatgcy cgcggaaca aaaactactg ggcgcccgga
 720

ggacccctcg acacctccaa gcccttcacc atggtgaccc agttcatcac ggacgacggc
 780

10 accagctcgg gcaagcttag ccagatcgtg cgctcctacg tgcaaaacgg ccagaaggtc
 840

15 gccagcgccg tgtccggcgg cgacagcatc accgtcgagg gctgctcgtc cgcctcagcc
 900

tacggcgggc tcgccaccat gggcaaggcc ctgagcagcg gcatgggtgt cgtgttcagc
 960

20 atttgaacg acaacagcca gtacatgaac tggctcgaca gcggcaacgc cggcccctgc
 1020

agcagcaccg agggcaaccc atccaacatc ctggccaaca accccaacac gcacgtcgtc
 1080

25 ttctccaaca tccgctgggg agacattggg tctactacga actcgactgc gccccgccc
 1140

30 ccgcctgcgt ccagcagcag gttttcgact acacggagga gctcgacgac ttcgagcagc
 1200

ccgagctgca cgcagactca ctgggggcag tgcggtggca ttgggtacag cgggtgcaag
 1260

35 acgtgcacgt cgggcactac gtgccagtat agcaacgact actactcgca atgcctttaa
 1320

40 <210> 26
 <211> 439
 <212> PRT
 <213> Artificial Sequence

45 <220>
 <223> 240H12
 <400> 26

50 Gln Gln Pro Gly Thr Ser Thr Pro Glu Val His Pro Lys Leu Thr Thr
 1 5 10 15

Tyr Lys Cys Thr Thr Ala Gly Gly Cys Val Ala Gln Asp Thr Ser Val
 20 25 30

55 Val Leu Asp Trp Gly Tyr His Trp Ile His Thr Val Asp Gly Tyr Thr
 35 40 45

	Ser	Cys	Thr	Thr	Ser	Ser	Gly	Val	Asp	Ser	Thr	Leu	Cys	Pro	Asp	Ala	
	50						55					60					
5	Ala	Thr	Cys	Ala	Lys	Asn	Cys	Val	Ile	Glu	Pro	Ala	Asn	Tyr	Thr	Ser	
	65					70					75					80	
10	Ala	Gly	Val	Thr	Thr	Ser	Gly	Asp	Ser	Leu	Thr	Met	Tyr	Gln	Tyr	Val	
					85					90					95		
15	Gln	Ser	Asn	Gly	Val	Tyr	Thr	Asn	Ala	Ser	Pro	Arg	Ala	Tyr	Leu	Leu	
				100					105					110			
20	Gly	Ala	Asp	Gly	Asp	Tyr	Val	Met	Phe	Lys	Leu	Leu	Asn	Gln	Glu	Leu	
			115					120					125				
25	Ser	Phe	Asp	Val	Asp	Val	Ser	Thr	Leu	Pro	Cys	Gly	Glu	Asn	Ala	Ala	
	130						135					140					
30	Leu	Tyr	Phe	Ser	Glu	Met	Asp	Lys	Thr	Gly	Gly	Arg	Asn	Glu	His	Asn	
	145					150					155					160	
35	Thr	Gly	Gly	Ala	Lys	Tyr	Gly	Ser	Gly	Tyr	Cys	Asp	Ala	Gln	Cys	Pro	
					165					170					175		
40	Val	Gln	Thr	Trp	Asn	Asn	Gly	Thr	Leu	Asn	Thr	Ser	His	Gln	Gly	Ser	
				180					185					190			
45	Cys	Cys	Asn	Glu	Met	Asp	Ile	Leu	Glu	Ala	Asn	Ser	Lys	Ala	Glu	Ala	
			195					200					205				
50	Phe	Thr	Pro	His	Pro	Cys	Ile	Gly	Asp	Asn	Cys	Asp	Lys	Gly	Gly	Cys	
	210						215					220					
55	Gly	Phe	Asn	Ser	Tyr	Ala	Arg	Gly	Asn	Lys	Asn	Tyr	Trp	Ala	Pro	Gly	
	225					230					235					240	
60	Gly	Thr	Leu	Asp	Thr	Ser	Lys	Pro	Phe	Thr	Met	Val	Thr	Gln	Phe	Ile	
					245					250					255		
65	Thr	Asp	Asp	Gly	Thr	Ser	Ser	Gly	Lys	Leu	Ser	Gln	Ile	Val	Arg	Ser	
				260					265					270			
70	Tyr	Val	Gln	Asn	Gly	Gln	Lys	Val	Ala	Ser	Ala	Val	Ser	Gly	Gly	Asp	
			275					280					285				

5 Ser Ile Thr Val Glu Gly Cys Ser Ser Ala Ser Ala Tyr Gly Gly Leu
 290 295 300

10 Ala Thr Met Gly Lys Ala Leu Ser Ser Gly Met Val Leu Val Phe Ser
 305 310 315 320

15 Ile Trp Asn Asp Asn Ser Gln Tyr Met Asn Trp Leu Asp Ser Gly Asn
 325 330 335

20 Ala Gly Pro Cys Ser Ser Thr Glu Gly Asn Pro Ser Asn Ile Leu Ala
 340 345 350

25 Asn Asn Pro Asn Thr His Val Val Phe Ser Asn Ile Arg Trp Gly Asp
 355 360 365

30 Ile Gly Ser Thr Thr Asn Ser Thr Ala Pro Pro Pro Pro Pro Ala Ser
 370 375 380

35 Ser Thr Thr Phe Ser Thr Thr Arg Arg Ser Ser Thr Thr Ser Ser Ser
 385 390 395 400

40 Pro Ser Cys Thr Gln Thr His Trp Gly Gln Cys Gly Gly Ile Gly Tyr
 405 410 415

45 Ser Gly Cys Lys Thr Cys Thr Ser Gly Thr Thr Cys Gln Tyr Ser Asn
 420 425 430

50 Asp Tyr Tyr Ser Gln Cys Leu
 435

55 <210> 27
 <211> 1131
 <212> DNA
 <213> Chaetomium globosum

60 <400> 27
 cagcaaccgg gttcgggtgac ccccgaggtc catcccaagt tgcccacctg gcagtgtaca

65 aagtccgggg ggtgcgtgga gcaggacacc tcggtggtcc ttgactggaa ctaccgctgg
 120

70 ttccacacct cggacaacac cacctcgtgc accacgtcct cgggcatcga ctcgacgctc
 180

75 tgccctgacg cagcgacctg tgccgagaac tgcgtcgtgg agggcaccga ctacaccagc
 240

tcgggcatcg agacctcggg cagcagcctc accctgaggg agttcgtcaa ggacagcgag
 300

5 ggcaagacca acagcgtctc tcctcggggc tatctcctgg gcgcgcgacgg tgactacgtg
 360

atgttcaagc tcctcaacca ggagctgagc ttcgacgtcg acgtctctac gctgccgtgt
 420

10 ggagagaacg cagcgtctta cttctctgag atggacaaga cggggggccg gaacgagcac
 480

15 aacacggggc gtgccaagta cgggagcggc tactgcgatg ctcagtggcc cgtccagaca
 540

tggaacaacg gcaccctcaa cactagccac cagggctcgt gctgcaacga gatggatata
 600

20 ctggaggcca actcgaaggc ggaagccttc acccctcacc cctgcacggc cgacaactgc
 660

gacaagggcg gttgcggctt caactcgtat gcgcgcggca acaaaaacta ctgggcgccc
 720

25 ggaggcaccg tcgacacctc caagcccttc accatgggtga cccagttcat cacggacgac
 780

30 ggcaccagct cgggcaagct tagccagatc gtgcgctcct acgtgcaaaa cggccagaag
 840

gtcgccagcg ccgtgtccgg cggcgacagc atcacgctcg agggctgctc gtctccgac
 900

35 gcctacggcg gcctcgtcgg tatgggagag gccctggggc ggggcatggg gctcgccatg
 960

agcatttgga acgacgccag cggcttcatg aactggctcg acagcggcga caacggcccc
 1020

40 tgcaacgaga ccgagggcga cccagccaac atcctggcca accaccccga ttgcagggtc
 1080

45 gtctgtcca acatccgctg gggagacatt gactctactg ttcagctcta a
 1131

50 <210> 28
 <211> 376
 <212> PRT
 <213> Chaetomium globosum

<400> 28

55 Gln Gln Pro Gly Ser Val Thr Pro Glu Val His Pro Lys Leu Pro Thr
 1 5 10 15

Trp Gln Cys Thr Lys Ser Gly Gly Cys Val Glu Gln Asp Thr Ser Val

	20							25					30				
5	Val	Leu	Asp	Trp	Asn	Tyr	Arg	Trp	Phe	His	Thr	Ser	Asp	Asn	Thr	Thr	
			35					40					45				
10	Ser	Cys	Thr	Thr	Ser	Ser	Gly	Ile	Asp	Ser	Thr	Leu	Cys	Pro	Asp	Ala	
		50					55					60					
15	Ala	Thr	Cys	Ala	Glu	Asn	Cys	Val	Val	Glu	Gly	Thr	Asp	Tyr	Thr	Ser	
	65					70					75					80	
20	Ser	Gly	Ile	Glu	Thr	Ser	Gly	Ser	Ser	Leu	Thr	Leu	Arg	Gln	Phe	Val	
					85					90					95		
25	Lys	Asp	Ser	Glu	Gly	Lys	Thr	Asn	Ser	Val	Ser	Pro	Arg	Ala	Tyr	Leu	
				100					105					110			
30	Leu	Gly	Ala	Asp	Gly	Asp	Tyr	Val	Met	Phe	Lys	Leu	Leu	Asn	Gln	Glu	
			115					120					125				
35	Leu	Ser	Phe	Asp	Val	Asp	Val	Ser	Thr	Leu	Pro	Cys	Gly	Glu	Asn	Ala	
		130					135					140					
40	Ala	Leu	Tyr	Phe	Ser	Glu	Met	Asp	Lys	Thr	Gly	Gly	Arg	Asn	Glu	His	
	145					150					155					160	
45	Asn	Thr	Gly	Gly	Ala	Lys	Tyr	Gly	Ser	Gly	Tyr	Cys	Asp	Ala	Gln	Cys	
					165					170					175		
50	Pro	Val	Gln	Thr	Trp	Asn	Asn	Gly	Thr	Leu	Asn	Thr	Ser	His	Gln	Gly	
				180					185					190			
55	Ser	Cys	Cys	Asn	Glu	Met	Asp	Ile	Leu	Glu	Ala	Asn	Ser	Lys	Ala	Glu	
			195					200					205				
60	Ala	Phe	Thr	Pro	His	Pro	Cys	Ile	Gly	Asp	Asn	Cys	Asp	Lys	Gly	Gly	
		210					215					220					
65	Cys	Gly	Phe	Asn	Ser	Tyr	Ala	Arg	Gly	Asn	Lys	Asn	Tyr	Trp	Ala	Pro	
	225					230					235					240	
70	Gly	Gly	Thr	Leu	Asp	Thr	Ser	Lys	Pro	Phe	Thr	Met	Val	Thr	Gln	Phe	
					245					250					255		

Ile Thr Asp Asp Gly Thr Ser Ser Gly Lys Leu Ser Gln Ile Val Arg
260 265 270

5 Ser Tyr Val Gln Asn Gly Gln Lys Val Ala Ser Ala Val Ser Gly Gly
275 280 285

10 Asp Ser Ile Thr Val Glu Gly Cys Ser Ser Ser Asp Ala Tyr Gly Gly
290 295 300

15 Leu Val Gly Met Gly Glu Ala Leu Gly Arg Gly Met Val Leu Ala Met
305 310 315 320

Ser Ile Trp Asn Asp Ala Ser Gly Phe Met Asn Trp Leu Asp Ser Gly
325 330 335

20 Asp Asn Gly Pro Cys Asn Glu Thr Glu Gly Asp Pro Ala Asn Ile Leu
340 345 350

25 Ala Asn His Pro Asp Ser Gln Val Val Leu Ser Asn Ile Arg Trp Gly
355 360 365

30 Asp Ile Asp Ser Thr Val Gln Leu
370 375

35 <210> 29
<211> 1323
<212> DNA
<213> *Aspergillus fumigatus*

40 <400> 29
cagcaaccgg ccgcgagttc tgctggtaac cccaagttga caacctacaa gtgtacaact
60

gctgggggggt gcgtggccca ggacacctcg gtggtccttg actggggcta ccaactggatt
120

45 cacacggtcg atgggtacac atcgtgcacc acatcgctcg gcgtcgacag cacgctctgc
180

cctgacgcgg cgacctgtgc gaagaactgc gtgatcgagc cggccaacta caccagcgcc
240

50 ggcgtcacga cctcgggcga cagcctcacc atgtaccagt acgttcagag caacggcgtc
300

55 tacaccaacg cctctcctcg gctgtatctc ctgggccccg acaagaacta cgtgatgctg
360

aagctcctag gccaggagct gaccttcgac gtcgacctct ctacactgcc gtgtggagag
420

aacggcgccc tctacctgtc tgaaatgagc gccaccgggg gccgcaacga atataacacg
 480
 5 ggcggtgccg agtacgggag cggctactgc gatgctcagt gccccgtcat cgcttgaag
 540
 aacggcaccc tcaacactag cggcgcaagc tactgctgca acgagatgga tatcctggag
 600
 10 gccaaactcga gggcgaattc gtacaccct caccctgca gtgccacgga ctgcgacaag
 660
 ggcggttgcg gcttcaacc ctatgctctc ggccaaaaa gctactgggg ccccgaggcg
 720
 15 accgttgaca cctccaagcc cttcaccatc accaccagc tcatcacgaa cgacggcacc
 780
 20 accaccggca ccctttccga aatccggcgc cagtacatgc aaaacggcaa ggtgatcgcc
 840
 aatgccgttt cctccactgg cgtcaactcc atcaccgagg actggtgcac gtccgtcgac
 900
 25 ggctcggccg ccacctttgg cggcctcacc accatgggca aggccctggg ccgcggcatg
 960
 gtgctcatct tcagcatttg gaacgacgcc agcggcttta tgaactggct cgacagcggc
 1020
 30 aacgccggcc cctgcagcag caccgagggc aaccagact tgatcaaggc ccagaacccc
 1080
 acgacgcacg tcgtcttctc caacatccgc tggggagaca ttgggtctac tttcaagggc
 1140
 35 tctgatggct cggtgacgac gacgacgtc actacatcga ccaagaccac gacttcgacc
 1200
 40 gcgcccgggc caacgcagac tcactatggg cagtgcggtg gccaaaggtg gactgggccc
 1260
 acggcttgcg catcgcccta cacgtgccag gttctgaacc cgtggtaact gcaatgcctt
 1320
 45 taa
 1323
 50 <210> 30
 <211> 440
 <212> PRT
 <213> *Aspergillus fumigatus*
 55 <400> 30

Gln Gln Pro Ala Ala Ser Ser Ala Gly Asn Pro Lys Leu Thr Thr Tyr
 1 5 10 15

Lys Cys Thr Thr Ala Gly Gly Cys Val Ala Gln Asp Thr Ser Val Val
 20 25 30
 5
 Leu Asp Trp Gly Tyr His Trp Ile His Thr Val Asp Gly Tyr Thr Ser
 35 40 45
 10
 Cys Thr Thr Ser Ser Gly Val Asp Ser Thr Leu Cys Pro Asp Ala Ala
 50 55 60
 15
 Thr Cys Ala Lys Asn Cys Val Ile Glu Pro Ala Asn Tyr Thr Ser Ala
 65 70 75 80
 20
 Gly Val Thr Thr Ser Gly Asp Ser Leu Thr Met Tyr Gln Tyr Val Gln
 85 90 95
 Ser Asn Gly Val Tyr Thr Asn Ala Ser Pro Arg Leu Tyr Leu Leu Gly
 100 105 110
 25
 Pro Asp Lys Asn Tyr Val Met Leu Lys Leu Leu Gly Gln Glu Leu Thr
 115 120 125
 30
 Phe Asp Val Asp Leu Ser Thr Leu Pro Cys Gly Glu Asn Gly Ala Leu
 130 135 140
 35
 Tyr Leu Ser Glu Met Ser Ala Thr Gly Gly Arg Asn Glu Tyr Asn Thr
 145 150 155 160
 40
 Gly Gly Ala Glu Tyr Gly Ser Gly Tyr Cys Asp Ala Gln Cys Pro Val
 165 170 175
 Ile Ala Trp Lys Asn Gly Thr Leu Asn Thr Ser Gly Ala Ser Tyr Cys
 180 185 190
 45
 Cys Asn Glu Met Asp Ile Leu Glu Ala Asn Ser Arg Ala Asn Ser Tyr
 195 200 205
 50
 Thr Pro His Pro Cys Ser Ala Thr Asp Cys Asp Lys Gly Gly Cys Gly
 210 215 220
 55
 Phe Asn Pro Tyr Ala Leu Gly Gln Lys Ser Tyr Trp Gly Pro Gly Gly
 225 230 235 240
 Thr Val Asp Thr Ser Lys Pro Phe Thr Ile Thr Thr Gln Phe Ile Thr
 245 250 255

5

10

15

20

25

30

35

40

45

50

Patentkrav

- 1.** Isoleret eller oprenset polypeptid **kendetegnet ved, at** det har en forbedret endoglucanase-aktivitet i forhold til endoglucanase-aktiviteten af
- 5 referenceproteinet EG1, hvor nævnte polypeptid er er valgt fra gruppen bestående af:
- i) en aminosyresekvens valgt fra SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24 og SEQ ID NO: 26;
 - 10 ii) en aminosyresekvens med, i forhold til sekvensen SEQ ID NO: 6, SEQ ID NO: 10, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24 eller SEQ ID NO: 26, en procentmæssig identitet på mindst 90 %, fortrinsvis 95 %, 98 % eller 99 %, eller en aminosyresekvens med, i forhold til sekvensen SEQ ID NO: 8, SEQ ID NO: 12, SEQ ID NO: 14 eller
 - 15 SEQ ID NO: 16, en procentmæssig identitet på mindst 98 %.
- 2.** Oprensset eller isoleret nukleinsyre, **kendetegnet ved, at** den koder for mindst et polypeptid ifølge krav 1.
- 20 **3.** Oprensset eller isoleret nukleinsyre ifølge krav 2 valgt fra de følgende sekvenser: SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11; SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23 og SEQ ID NO: 25.
- 25 **4.** Vektor **kendetegnet ved, at** den omfatter en nukleinsyre ifølge et af kravene 2 eller 3.
- 5.** Isoleret værtscelle **kendetegnet ved, at** den omfatter nukleinsyren ifølge et af kravene 2 eller 3 eller vektoren ifølge krav 4.
- 30 **6.** Isoleret værtscelle ifølge krav 5, **kendetegnet ved, at** den er valgt fra *Trichoderma*, *Aspergillus*, *Neurospora*, *Humicola*, *Penicillium*, *Fusarium*,

Thermomonospora, Myceliophthora, Chrysosporium, Bacillus, Pseudomonas, Escherichia, Clostridium, Cellulomonas, Streptomyces, Yarrowia, Pichia og Saccharomyces.

- 5 **7.** Isoleret værtscelle ifølge krav 5 eller 6, **kendetegnet ved, at** den er valgt fra *Trichoderma reesei, Trichoderma viridae, Trichoderma koningii, Aspergillus niger, Aspergillus nidulans, Aspergillus wentii, Aspergillus oryzae, Aspergillus phoenicis, Neurospora crassa, Humicola grisea, Myceliophthora thermopila, Chrysosporium lucknowense, Penicillium pinophilum, Penicillium oxalicum, Escherichia coli,*
- 10 *Clostridium acetobutylicum, Clostridium saccharolyticum, Clostridium benjerinckii, Clostridium butylicum, Pichia pastoris, Yarrowia lipolityca, Saccharomyces cerevisiae.*

- 8.** Anvendelse af nævnte polypeptid ifølge krav 1 til hydrolysen af cellulose.
- 15

9. Anvendelse af nævnte polypeptid ifølge krav 1 til fremstillingen af biobrændstof.

- 10.** Enzymatisk sammensætning i stand til at virke på lignocellulose biomasse,
- 20 hvor nævnte enzymatiske sammensætning fremstilles af filamentøse svampe og omfatter mindst et polypeptid ifølge krav 1.

- 11.** Fremgangsmåde til fremstilling af biobrændstof fra biomasse, **kendetegnet ved, at** den omfatter de følgende successive trin:

- 25 - opslæmning i vandig fase af biomassen, som skal hydrolyseres;
- hydrolyse under tilstedeværelse af en enzymatisk sammensætning ifølge krav 10 af lignocellulose biomassen for at fremstille et hydrolysat indeholdende glukose;
- fermentering af hydrolysatets glukose for at fremstille en
- 30 fermentationsmost;
- separation af biobrændstoffet fra fermentationsmosten.

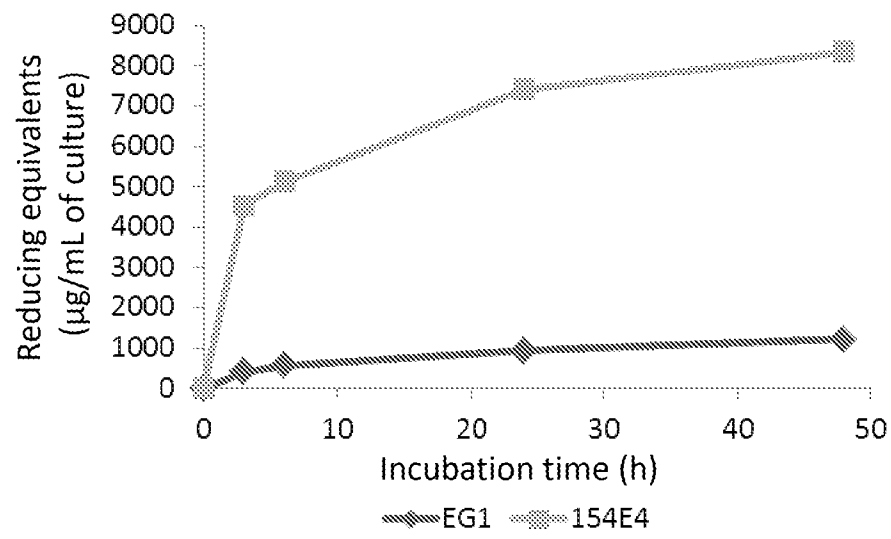


FIGURE 1

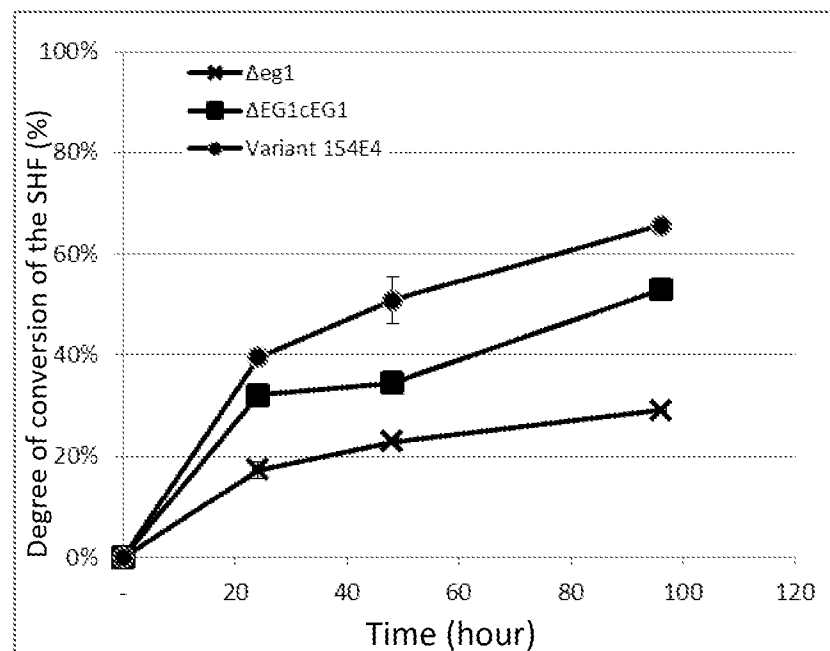


FIGURE 2

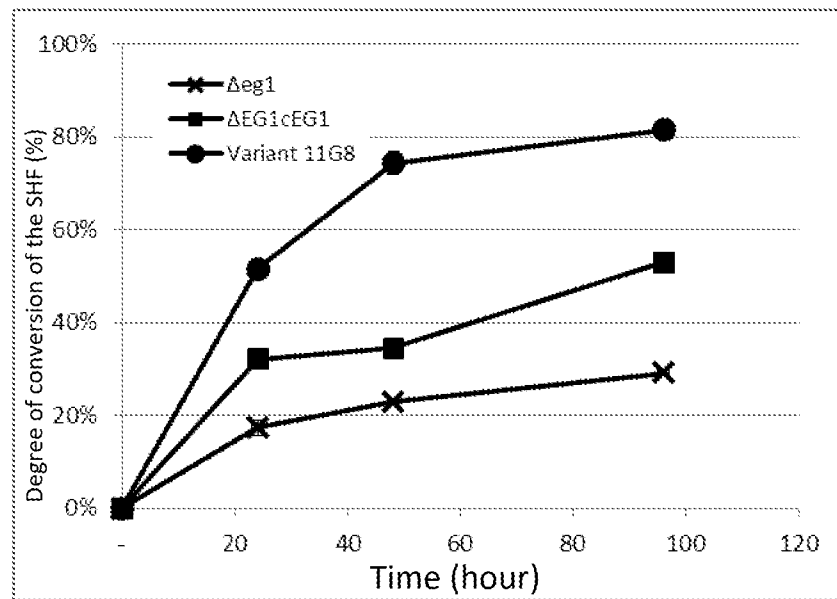


FIGURE 3

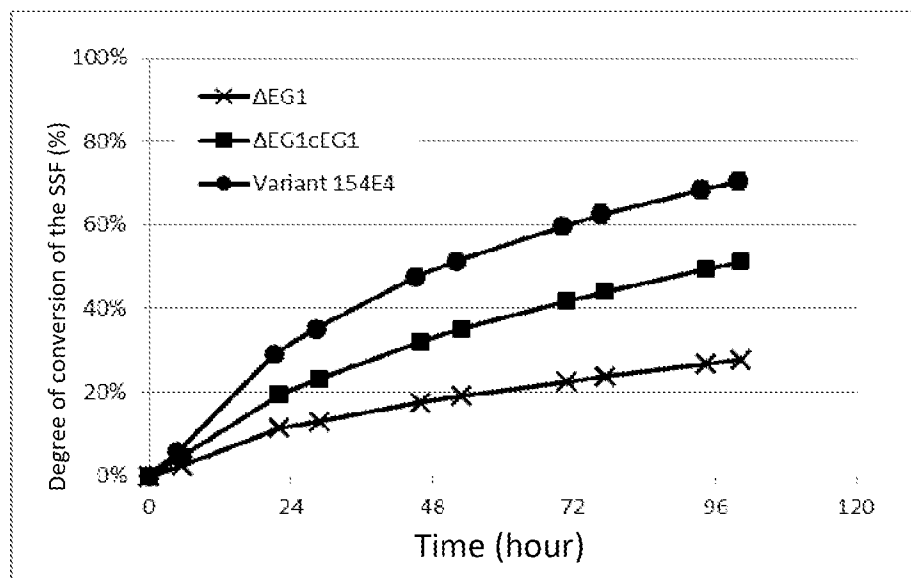


FIGURE 4

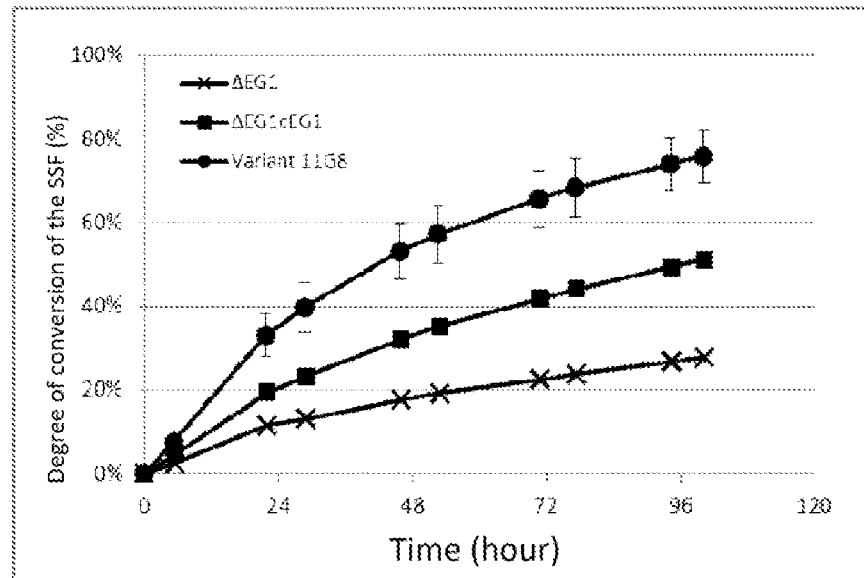


FIGURE 5