

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
17 January 2008 (17.01.2008)

PCT

(10) International Publication Number
WO 2008/008070 A2

(51) International Patent Classification:

C12P 7/10 (2006.01)

(21) International Application Number:

PCT/US2006/027347

(22) International Filing Date: 6 October 2006 (06.10.2006)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

11/487,547 13 July 2006 (13.07.2006) US

(71) Applicant (for all designated States except US): **DYADIC INTERNATIONAL, INC.** [US/US]; 140 Intracoastal Pointe Drive, Jupiter, Florida 33477-5094 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **GUSAKOV, Alexander V.** [RU/US]; C/o Dyadic International, Inc., 140 Intracoastal Pointe Drive, Suite 404, Jupiter, Florida 33477-5094 (US). **SALANOVICH, Tatyana N.** [RU/US]; C/o Dyadic International, Inc., 140 Intracoastal Pointe Drive, Suite 404, Jupiter, FL 33477-5094 (US). **ANTONOV, Alexey I.** [RU/US]; C/o Dyadic International, Inc., 140 Intracoastal Pointe Drive, Suite 404, Jupiter, Florida 33477-5094 (US). **USTINOV, Boris B.** [RU/US]; C/o Dyadic International, Inc., 140 Intracoastal Pointe Drive, Suite 404, Jupiter, Florida 33477-5094

(US). **OKUNEV, Oleg N.** [RU/US]; C/o Dyadic International, Inc., 140 Intracoastal Pointe Drive, Suite 404, Jupiter, FL 33477-5094 (US). **BURLINGAME, Richard** [US/US]; C/o Dyadic International, Inc., 140 Intracoastal Pointe Drive, Suite 404, Jupiter, FL 33477-5094 (US). **EMALFARB, Mark** [US/US]; C/o Dyadic International, Inc., 140 Intracoastal Pointe Drive, Suite 404, Jupiter, Florida 33477-5094 (US). **BAEZ, Marco** [MX/US]; C/o Dyadic International, Inc., 140 Intracoastal Pointe Drive, Suite 404, Jupiter, Florida 33477-5094 (US). **SINITSYN, Arkady P.** [RU/US]; C/o Dyadic International, Inc., 140 Intracoastal Pointe Drive, Suite 404, Jupiter, Florida 33477-5094 (US).

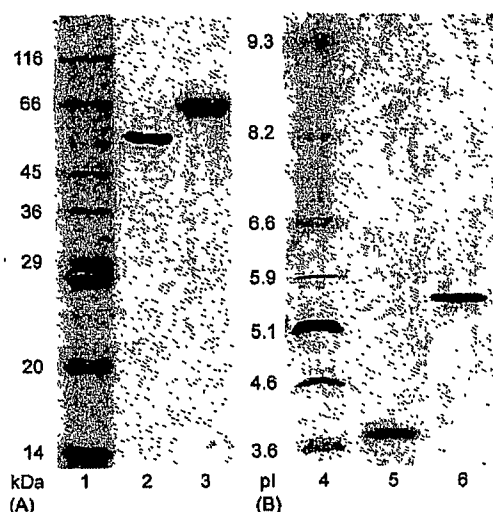
(74) Agent: **ENG, JR., Joseph D.**; Morgan & Finnegan, LLP, 3 World Financial Center, New York, ny 10281-2101 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM,

[Continued on next page]

(54) Title: CONSTRUCTION OF HIGHLY EFFICIENT CELLULASE COMPOSITIONS FOR ENZYMATIC HYDROLYSIS OF CELLULOSE



(57) Abstract: This invention provides novel enzyme compositions using newly identified and isolated *C. lucknowense* enzymes, including CBH Ib, CBH Iib, EG II, EG VI, β -glucosidase, and xylanase II in conjunction with previously identified enzymes CBH Ia, CBH Iia (previously described as Endo 43), and EG V. These enzyme compositions demonstrate an extremely high ability to convert lignocellulosic biomass (e.g., Avicel, cotton, Douglas fir wood pretreated by organosolv) to glucose. CBH Ia and Iib, which both have a cellulose-binding module (CBM) displayed a pronounced synergism with three major endoglucanases (EG II, EG V, EG VI) from the same fungus in hydrolysis of cotton as well as a strong synergy with each other. The enzyme compositions are effective in hydrolysis of the lignocellulosic biomass.



ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

Published:

- *without international search report and to be republished upon receipt of that report*

CONSTRUCTION OF HIGHLY EFFICIENT CELLULASE COMPOSITIONS FOR ENZYMATIC HYDROLYSIS OF CELLULOSE

[0001] This application is a continuation-in-part of U.S. Patent Application 10/394,568, filed March 21, 2003 which claims the benefit of U.S. Patent Application 09/548,938 (now US Patent 6,573,086), filed April 13, 2000, which is a continuation-in-part of international application PCT/NL99/00618, filed Oct. 6, 1999, which is a continuation-in-part of international application PCT/EP98/06496, filed Oct. 6, 1998. This application is also a continuation-in-part application of U.S. Patent Application 09/284,152, filed on April 8, 1999 which is a continuation-in-part of 08/731,170 filed October 10, 1996. All prior applications to which priority is claimed are hereby incorporated by reference in their entirety.

FIELD OF THE INVENTION

[0002] This invention relates to compositions and methods for producing bioenergy or other value-added products from lignocellulosic biomass or cellulosic materials. In particular, the invention provides enzyme compositions capable of converting a variety of cellulosic substrates or lignocellulosic biomass into a fermentable sugar. The invention also provides methods for using such enzyme compositions.

INTRODUCTION

[0003] Bioconversion of renewable lignocellulosic biomass to a fermentable sugar that is subsequently fermented to produce alcohol (e.g., ethanol) as an alternative to liquid fuels has attracted an intensive attention of researchers since 1970s, when the oil crisis broke out because of decreasing the output of petroleum by OPEC (Bungay H.R., "Energy: the biomass options". NY: Wiley; 1981; Olsson L, Hahn-Hägerdal B. "Fermentation of lignocellulosic hydrolysates for ethanol production", *Enzyme Microb*

Technol 1996;18:312-31; Zaldivar J, Nielsen J, Olsson L. "Fuel ethanol production from lignocellulose: a challenge for metabolic engineering and process integration", *Appl Microbiol Biotechnol* 2001;56:17-34; Galbe M, Zacchi G., "A review of the production of ethanol from softwood", *Appl Microbiol Biotechnol* 2002; 59:618-28). Ethanol has been widely used as a 10% blend to gasoline in the USA or as a neat fuel for vehicles in Brazil in the last two decades. The importance of fuel bioethanol will increase in parallel with skyrocketing prices for oil and gradual depletion of its sources. Additionally, fermentable sugars are being used to produce plastics, polymers and other biobased products and this industry is expected to grow substantially therefore increasing the demand for abundant low cost fermentable sugars which can be used as a feed stock in lieu of petroleum based feedstocks (e.g. see article "The Rise Of Industrial Biotech" published in Forbes July 24, 2006)

[0004] The major polysaccharides comprising different lignocellulosic residues, which may be considered as a potential renewable feedstock, are cellulose and hemicelluloses (xylans). The enzymatic hydrolysis of these polysaccharides to soluble sugars, for example glucose, xylose, arabinose, galactose, mannose, and other hexoses and pentoses occurs under the action of different enzymes acting in concert. Endo-1,4- β -glucanases (EG) and exo-cellobiohydrolases (CBH) catalyze the hydrolysis of insoluble cellulose to celooligosaccharides (cellobiose as a main product), while β -glucosidases (BGL) convert the oligosaccharides to glucose. Xylanases together with other accessory enzymes (non-limiting examples of which include α -L-arabinofuranosidases, feruloyl and acetylxytan esterases, glucuronidases, and β -xylosidases) catalyze the hydrolysis of hemicelluloses.

[0005] Regardless of the type of cellulosic feedstock, the cost and hydrolytic efficiency of enzymes are major factors that restrict the commercialization of the biomass bioconversion processes. The production costs of microbially produced enzymes are tightly connected with a productivity of the enzyme-producing strain and the final activity yield in the fermentation broth. The hydrolytic efficiency of a multienzyme complex in the process of lignocellulose saccharification depends both on properties of individual enzymes, the synergies between them, and their ratio in the multienzyme cocktail.

[0006] *Chrysosporium lucknowense* is a fungus that is known to produce a wide variety of cellulases, hemicellulases, and possibly other accessory enzymes. *C. lucknowense* also secretes at least five different endoglucanases, the EG II (51 kDa, Cel5A) being the most active. Moreover, *C. lucknowense* mutant strains (including UV18-25) have been developed to produce enzymes for textile, pulp and paper, detergent and other applications, but not for the enzymatic saccharification of cellulose; these strains can also be used for a high-level production of homologous and heterologous proteins. The best *C. lucknowense* mutant strains secrete at least 50-80 g l⁻¹ of extracellular protein in low viscosity fermentations. The full fungal genome of the *C. lucknowense* has been sequenced in 2005 (see http://WorldWideWeb.dyadic-group.com/wt/dyad/pr_1115654417), and now the genome annotation is being carried out.

[0007] The crude *C. lucknowense* multienzyme complex demonstrates modest results in cellulose saccharification, with only a fraction of the cellulose being converted to glucose under the conditions tested. Two cellobiohydrolases of *C. lucknowense*, belonging to families 7 and 6 of glycoside hydrolases: CBH Ia (Cel7A) and CBH IIa (Cel6A), have been previously isolated and studied. CBH Ia was previously referred to as CBH I, 70(60) kD protein in US Patent No. 6,573,086. CBH Ia exists in the culture broth as a full size enzyme (observed molecular mass 65 kDa, SDS-PAGE data), consisting of a core catalytic domain and cellulose-binding module (CBM) connected by a flexible peptide linker, and its truncated form (52 kDa), representing the enzyme catalytic domain. CBH I (Cel7A) of *C. lucknowense* appears to be slightly less effective in hydrolysis of crystalline cellulose but more thermostable than the CBH I of *T. reesei*. CBH IIa was previously thought to be an endoglucanase and has been referred to as 43 kD Endo and EG6. See, e.g., US Patent No. 6,573,086. CBH IIa (43 kDa) has no CBM, i.e. its molecule contains only the catalytic domain.

[0008] In spite of the continued research of the last few decades to understand enzymatic lignocellulosic biomass degradation and cellulase production, it remains desirable to discover or to engineer new highly active cellulases and hemicellulases. It would also be highly desirable to construct highly efficient enzyme compositions capable of performing rapid and efficient biodegradation of lignocellulosic materials.

SUMMARY OF THE INVENTION

[0009] This invention provides several newly identified and isolated enzymes from *C. lucknowense*. The new enzymes include two new cellobiohydrolases (CBH Ib and Iib, or Cel7B and Cel6B), an endoglucanase (EG VI), (not to be confused with CBH IIa, which was previously referred to as EG 6)a β -glucosidase (BGL), and a xylanase (Xyl II). The CBH Iib has a high activity against Avicel and cotton and displayed a pronounced synergism with other *C. lucknowense* cellulases. Using these new enzymes, this invention provides highly effective enzyme compositions for cellulose hydrolysis.

[0010] One object of this invention is to provide an enzyme formulation that includes at least one isolated cellobiohydrolase obtained from *C. lucknowense*. The isolated cellobiohydrolase may be either CBH Ib and Iib. The enzyme formulation may optionally contain an endoglucanase and/or a β -glucosidase. Furthermore, the enzyme formulation may optionally contain a hemicellulase.

[0011] Another object of this invention is to provide a method for producing glucose from cellulose. The method includes producing an enzyme formulation that contains at least one isolated cellobiohydrolase obtained from *C. lucknowense*, which can be CBH Ib or Iib. Optionally, the enzyme formulation may contain an endoglucanase and/or a β -glucosidase. The enzyme formulation is applied to cellulose to form glucose.

[0012] Yet another aspect of this invention is to provide a method of producing ethanol. The method includes providing an enzyme formulation that contains at least one isolated cellobiohydrolase obtained from *C. lucknowense*, which can be CBH Ib or Iib. The enzyme formulation optionally may contain an endoglucanase and/or a β -glucosidase. Furthermore, the enzyme formulation may optionally contain a hemicellulase. The method further includes applying the enzyme formulation to cellulose to produce glucose and subsequently fermenting the glucose to produce ethanol.

[0013] This invention also provides a method of producing energy from ethanol. The method includes providing an enzyme formulation that contains at least one

isolated cellobiohydrolase obtained from *C. lucknowense*, which can be CBH Ib or IIb. The enzyme formulation optionally may contain an endoglucanase and/or a β -glucosidase. Furthermore, the enzyme formulation may optionally contain a hemicellulase. The method further includes applying the enzyme formulation to cellulose to produce glucose, fermenting the glucose to produce ethanol, and combusting said ethanol to produce energy.

[0014] Another aspect of this invention is to provide a mutant *Chrysosporium lucknowense* strain capable of expressing at least one cellobiohydrolase and at least one endo-1,4- β -glucanase at higher levels than the corresponding non-mutant strain under the same conditions. The cellobiohydrolase is selected from the group consisting of CBH Ia, CBH IIa, CBH Ib, and CBH IIb; and the endo-1,4- β -glucanase is selected from the group consisting of EG II, EG V, and EG VI.

[0015] Yet another aspect of this invention is to provide proteins exhibiting at least 65% amino acid identity as determined by the BLAST algorithm with the CBH Ib, CBH IIb, EG VI, BGL, and Xyl II amino acid sequences of SEQ ID NOs. 2, 4, 16, 12, and 18, respectively, or a part thereof having at least 20 contiguous amino acids. This invention also contemplates the corresponding nucleic acid sequences that encode such a protein.

[0016] One aspect of this invention provides an enzyme formulation comprising at least one enzyme selected from the group consisting of CBH Ib, CBH IIb, EG II, EG VI, BGL, and Xyl II.

[0017] Another aspect of this invention provides a method of producing fermentable sugars from lignocellulosic material. The method comprises (a) providing an enzyme formulation comprising at least one enzyme selected from the group consisting of CBH Ib, CBH IIb, EG II, EG VI, BGL, and Xyl II; and (b) applying the enzyme formulation to lignocellulosic material to produce fermentable sugars.

[0018] The invention also provides a method of producing a fermentation product or a starting material for a fermentation product from a fermentable sugar. This method comprises (a) providing an enzyme formulation, wherein the enzyme formulation contains at least one enzyme selected from the group consisting of CBH Ib, CBH IIb, EG II, EG VI, BGL, and Xyl II; (b) applying the enzyme formulation to

lignocellulosic material to produce a fermentable sugar; and (c) fermenting said fermentable sugar to produce a fermentation product.

[0019] In another aspect, the invention provides a method of producing energy from a fermentable sugar. The method comprises (a) providing an enzyme formulation, wherein the enzyme formulation comprises at least one enzyme selected from the group consisting of CBH Ib, CBH IIb, EG II, EG VI, BGL, and Xyl II; (b) applying the enzyme formulation to lignocellulosic material to produce a fermentable sugar; (c) fermenting the fermentable sugar to produce a combustible fermentation product; and (d) combusting said combustible fermentation product to produce energy.

[0020] One object of the invention is provide a mutant *Chrysosporium lucknowense* strain capable of expressing at least one cellobiohydrolase and at least one endo-1,4- β -glucanase at higher levels than the corresponding non-mutant strain under the same conditions. The cellobiohydrolase is selected from the group consisting of CBH Ia, CBH Ib, CBH IIa and CBH IIb; and the endo-1,4- β -glucanase is selected from the group consisting of EG II, EG V, and EG VI.

[0021] The invention also provides a protein exhibiting at least 65% amino acid identity as determined by the BLAST algorithm with the CBH Ib, IIb, EG VI, BGL, Xyl II amino acid sequences as defined herein or a part thereof having at least 20 contiguous amino acids.

[0022] Another aspect of this invention provides a nucleic acid sequence having at least 80% homology with the nucleic acid sequence encoding CBH Ib, CBH IIb, EG II, EG VI, BGL, or Xyl II, as defined herein.

[0023] The invention also provides a method for degrading a lignocellulosic material to fermentable sugars. The method includes contacting the lignocellulosic material with an effective amount of a multi-enzyme product derived from a microorganism, to produce at least one fermentable sugar. At least one enzyme in the multi-enzyme product is selected from the group consisting of CBH Ia, CBH Ib, CBH IIa, CBH IIb, EG II, EG V, EG VI, BGL, and Xyl II.

[0024] In another aspect, the invention provides a microorganism or plant capable of expressing one or more of an enzyme selected from the group consisting of CBH Ia, CBH Ib, CBH IIa, CBH IIb, EG II, EG V, EG VI, BGL, and Xyl II.

BRIEF DESCRIPTION OF THE DRAWINGS

[0025] **FIG. 1:** SDS/PAGE (A) and isoelectrofocusing (B) of purified cellobiohydrolases from *C. lucknowense*. Lanes: 1, markers with different molecular masses; 2 and 5, CBH Ib; 3 and 6, CBH IIb; 4, markers with different pI.

[0026] **FIG. 2:** Progress kinetics of Avicel (5 mg ml⁻¹) hydrolysis by purified cellobiohydrolases (0.1 mg ml⁻¹) in the presence of purified *A. japonicus* BGL (0.5 U ml⁻¹), 40°C, pH 5.0.

[0027] **FIG. 3:** Synergism between CBH IIb and other *C. lucknowense* purified enzymes during hydrolysis of cotton cellulose (5 mg ml⁻¹) in the presence of purified *A. japonicus* BGL (0.5 U ml⁻¹), 40°C, pH 5.0. The CBH and EG concentration was 0.15 and 0.05 mg ml⁻¹, respectively. Experimental data for the pairs of enzymes are shown with open symbols (continuous curves); the theoretical sums of glucose concentrations obtained under the action of individual enzymes are shown with filled symbols (dotted lines).

[0028] **FIG. 4:** Progress kinetics of cotton (25 mg ml⁻¹) hydrolysis by combination #1 of purified *C. lucknowense* enzymes and NCE L-600, a commercial *C. lucknowense* multienzyme cellulase preparation at protein loading of 0.5 mg ml⁻¹, 50°C, pH 5.0 (see text and Table 4 for details).

[0029] **FIG. 5:** Progress kinetics of Avicel (50 mg ml⁻¹) hydrolysis by combination #1 of purified *C. lucknowense* enzymes and NCE-L, a commercial *C. lucknowense* multienzyme cellulase preparation at protein loading of 0.5 mg ml⁻¹, 50°C, pH 5.0 (see text and Table 4 for details).

[0030] **FIG. 6:** Progress kinetics of hydrolysis of pretreated Douglas fir wood (50 mg ml⁻¹) by combination #1 of purified *C. lucknowense* enzymes and NCE-L 600, a

commercial *C. lucknowense* at protein loading of 0.5 mg ml⁻¹, 50°C, pH 5.0 (see text and Table 4 for details).

[0031] **FIG. 7:** Progress kinetics of hydrolysis of pretreated Douglas fir wood (50 mg ml⁻¹) by different combinations of purified *C. lucknowense* enzymes at protein loading of 0.5 mg ml⁻¹, 50°C, pH 5.0 (see text and Table 5 for details).

[0032] **FIG. 8:** cbh2 gene encoding CBH IB.

[0033] **FIG. 9:** cbh4 gene encoding CBH IIb

[0034] **FIG. 10:** cbh1 gene encoding CBH Ia

[0035] **FIG. 11:** EG6 gene encoding CBH IIa

[0036] **FIG. 12:** eg2 gene encoding EG II

[0037] **FIG. 13:** bgl1 gene encoding BGL

[0038] **FIG. 14:** eg5 gene encoding EGV

[0039] **FIG. 15:** eg7 gene encoding EG VI

[0040] **FIG. 16:** xyl2 gene encoding Xyl II

DETAILED DESCRIPTION OF THE INVENTION

[0041] The present invention provides methods and compositions for the conversion of plant biomass to fermentable sugars that can be converted to useful products. The methods include methods for degrading lignocellulosic material using enzyme mixtures to liberate sugars. The compositions of the invention include enzyme combinations that break down lignocellulose. As used herein the terms "biomass" or "lignocellulosic material" includes materials containing cellulose and/or hemicellulose. Generally, these materials also contain xylan, lignin, protein, and carbohydrates, such as starch and sugar. Lignocellulose is generally found, for example, in the stems, leaves, hulls, husks, and cobs of plants or leaves, branches, and wood of trees. The process of converting a complex carbohydrate (such as starch, cellulose, or hemicellulose) into fermentable sugars is also referred to herein as "saccharification." Fermentable sugars,

as used herein, refers to simple sugars, such as glucose, xylose, arabinose, galactose, mannose, rhamnose, sucrose and fructose.

[0042] Biomass can include virgin biomass and/or non-virgin biomass such as agricultural biomass, commercial organics, construction and demolition debris, municipal solid waste, waste paper and yard waste. Common forms of biomass include trees, shrubs and grasses, wheat, wheat straw, sugar cane bagasse, corn, corn husks, corn kernel including fiber from kernels, products and by-products from milling of grains such as corn, wheat and barley (including wet milling and dry milling) as well as municipal solid waste, waste paper and yard waste. The biomass can also be, but is not limited to, herbaceous material, agricultural residues, forestry residues, municipal solid wastes, waste paper, and pulp and paper mill residues. "Agricultural biomass" includes branches, bushes, canes, corn and corn husks, energy crops, forests, fruits, flowers, grains, grasses, herbaceous crops, leaves, bark, needles, logs, roots, saplings, short rotation woody crops, shrubs, switch grasses, trees, vegetables, fruit peels, vines, sugar beet pulp, wheat midlings, oat hulls, and hard and soft woods (not including woods with deleterious materials). In addition, agricultural biomass includes organic waste materials generated from agricultural processes including farming and forestry activities, specifically including forestry wood waste. Agricultural biomass may be any of the aforestated singularly or in any combination or mixture thereof.

[0043] The fermentable sugars can be converted to useful value-added fermentation products, non-limiting examples of which include amino acids, vitamins, pharmaceuticals, animal feed supplements, specialty chemicals, chemical feedstocks, plastics, solvents, fuels, or other organic polymers, lactic acid, and ethanol, including fuel ethanol. Specific value-added products that may be produced by the methods of the invention include, but not limited to, biofuels (including ethanol and butanol); lactic acid; plastics; specialty chemicals; organic acids, including citric acid, succinic acid and maleic acid; solvents; animal feed supplements; pharmaceuticals; vitamins; amino acids, such as lysine, methionine, tryptophan, threonine, and aspartic acid; industrial enzymes, such as proteases, cellulases, amylases, glucanases, lactases, lipases, lyases, oxidoreductases, transferases and xylanases; and chemical feedstocks.

[0044] As used herein, a multi-enzyme product can be obtained from or derived from a microbial, plant, or other source or combination thereof, and will contain enzymes capable of degrading lignocellulosic material. Examples of enzymes comprising the multi-enzyme products of the invention include cellulases (such as cellobiohydrolases, endoglucanase, β -glucosidases, hemicellulases (such as xylanases, including endoxylanases, exoxylanase, and β -xylosidase), ligninases, amylases, α -arabinofuranosidases, α -glucuronidases, α -glucuronidases, arabinases, glucuronidases, proteases, esterases (including ferulic acid esterase and acetylxytan esterase), lipases, glucomannanases, and xylogluconases.

[0045] In some embodiments, the multi-enzyme product comprises a hemicellulase. Hemicellulose is a complex polymer, and its composition often varies widely from organism to organism, and from one tissue type to another. In general, a main component of hemicellulose is beta-1,4-linked xylose, a five carbon sugar. However, this xylose is often branched as beta-1,3 linkages, and can be substituted with linkages to arabinose, galactose, mannose, glucuronic acid, or by esterification to acetic acid. Hemicellulose can also contain glucan, which is a general term for beta-linked six carbon sugars. Those hemicelluloses include xyloglucan, glucomannan, and galactomannan.

[0046] The composition, nature of substitution, and degree of branching of hemicellulose is very different in dicotyledonous plants (dicots, i.e., plant whose seeds have two cotyledons or seed leaves such as lima beans, peanuts, almonds, peas, kidney beans) as compared to monocotyledonous plants (monocots; i.e., plants having a single cotyledon or seed leaf such as corn, wheat, rice, grasses, barley). In dicots, hemicellulose is comprised mainly of xyloglucans that are 1,4-beta-linked glucose chains with 1,6-beta-linked xylosyl side chains. In monocots, including most grain crops, the principal components of hemicellulose are heteroxylans. These are primarily comprised of 1,4-beta-linked xylose backbone polymers with 1,3-beta linkages to arabinose, galactose and mannose as well as xylose modified by ester-linked acetic acids. Also present are branched beta glucans comprised of 1,3- and 1,4-beta-linked glucosyl chains. In monocots, cellulose, heteroxylans and beta glucans are present in roughly equal amounts, each comprising about 15-25% of the dry matter of cell walls.

[0047] Hemicellulolytic enzymes, i.e. hemicellulases, include includes both exohydrolytic and endohydrolytic enzymes, such as xylanase, β -xylosidase and esterases, which actively cleave hemicellulosic material through hydrolysis. These xylanase and esterase enzymes cleave the xylan and acetyl side chains of xylan and the remaining xylo-oligomers are unsubstituted and can thus be hydrolysed with Pxylosidase only. In addition, several less known side activities have been found in enzyme preparations which hydrolyse hemicellulose. While the multi-enzyme product may contain many types of enzymes, mixtures comprising enzymes that increase or enhance sugar release from biomass are preferred, including hemicellulases. In one embodiment, the hemicullulase is a xylanase, an arabinofuranosidase, an acetyl xylan esterase, a glucuronidase, an endo-galactanase, a mannanase, an endo arabinase, an exo arabinase, an exo-galactanase, a ferulic acid esterase, a galactomannanase, a xylogluconase, or mixtures of any of these. In particular, the enzymes can include glucoamylase, β -xylosidase and/or β -glucosidase. The enzymes of the multi-enzyme product can be provided by a variety of sources. In one embodiment, the enzymes can be produced by growing microorganisms or plants which produce the enzymes naturally or by virtue of being genetically modified to express the enzyme or enzymes. In another embodiment, at least one enzyme of the multi-enzyme product is commercially available.

[0048] One embodiment of the present invention relates to an isolated enzyme for catalyzing the conversion of lignocellulosic material to fermentable sugars as described herein, a homologue thereof, and/or a fragment thereof. Also included in the invention are isolated nucleic acid molecules encoding any of such proteins, homologues or fragments thereof. According to the present invention, an isolated protein or polypeptide is a protein that has been removed from its natural milieu (i.e., that has been subject to human manipulation) and can include purified proteins, partially purified proteins, recombinantly produced proteins, and synthetically produced proteins, for example. As such, "isolated" does not reflect the extent to which the protein has been purified. Preferably, an isolated protein of the present invention is produced recombinantly. An isolated peptide can be produced synthetically (e.g., chemically, such as by peptide synthesis) or recombinantly. An isolated protein can also be provided as a crude fermentation product, or a protein preparation that has been partially purified

or purified (e.g., from a microorganism) using protein purification procedures known in the art. In addition, and solely by way of example, a protein referenced as being derived from or from a particular organism, such as a "*Chrysosporium lucknowense* cellulase and/or hemicellulase" refers to a cellulase and/or hemicellulase (generally including a homologue of a naturally occurring cellulose and/or hemicellulase) from a *Chrysosporium lucknowense* microorganism, or to a cellulase and/or hemicellulase that has been otherwise produced from the knowledge of the structure (e.g., sequence), and perhaps the function, of a naturally occurring cellulase and/or hemicellulase from *Chrysosporium lucknowense*. In other words, general reference to a *Chrysosporium lucknowense* cellulase and/or hemicellulase or a cellulase and/or hemicellulase derived from *Chrysosporium lucknowense* includes any cellulase and/or hemicellulase that has substantially similar structure and function of a naturally occurring cellulase and/or hemicellulase from *Chrysosporium lucknowense* or that is a biologically active (i.e., has biological activity) homologue of a naturally occurring cellulase and/or hemicellulase from *Chrysosporium lucknowense* as described in detail herein. As such, a *Chrysosporium lucknowense* cellulase and/or hemicellulase can include purified, partially purified, recombinant, mutated/modified and synthetic proteins. The same description applies to reference to other proteins or peptides described herein and to other microbial sources for such proteins or peptides.

[0049] One embodiment of the present invention relates to isolated nucleic acid molecules comprising, consisting essentially of, or consisting of nucleic acid sequences that encode any of the enzymes described herein, including a homologue or fragment of any of such enzymes, as well as nucleic acid sequences that are fully complementary thereto. In accordance with the present invention, an isolated nucleic acid molecule is a nucleic acid molecule that has been removed from its natural milieu (i.e., that has been subject to human manipulation), its natural milieu being the genome or chromosome in which the nucleic acid molecule is found in nature. As such, "isolated" does not necessarily reflect the extent to which the nucleic acid molecule has been purified, but indicates that the molecule does not include an entire genome or an entire chromosome in which the nucleic acid molecule is found in nature. An isolated nucleic acid molecule can include a gene. An isolated nucleic acid molecule that includes a gene is not a fragment of a chromosome that includes such gene, but rather includes the coding

region and regulatory regions associated with the gene, but no additional genes that are naturally found on the same chromosome. An isolated nucleic acid molecule can also include a specified nucleic acid sequence flanked by (i.e., at the 5' and/or the 3' end of the sequence) additional nucleic acids that do not normally flank the specified nucleic acid sequence in nature (i.e., heterologous sequences). Isolated nucleic acid molecule can include DNA, RNA (e.g., mRNA), or derivatives of either DNA or RNA (e.g., cDNA). Preferably, an isolated nucleic acid molecule of the present invention is produced using recombinant DNA technology (e.g., polymerase chain reaction (PCR) amplification, cloning) or chemical synthesis. A nucleic acid molecule homologue can be produced using a number of methods known to those skilled in the art (see, for example, Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Labs Press (1989)). For example, nucleic acid molecules can be modified using a variety of techniques including, but not limited to, classic mutagenesis techniques and recombinant DNA techniques, such as site-directed mutagenesis, chemical treatment of a nucleic acid molecule to induce mutations, restriction enzyme cleavage of a nucleic acid fragment, ligation of nucleic acid fragments, PCR amplification and/or mutagenesis of selected regions of a nucleic acid sequence, synthesis of oligonucleotide mixtures and ligation of mixture groups to "build" a mixture of nucleic acid molecules and combinations thereof. Nucleic acid molecule homologues can be selected from a mixture of modified nucleic acids by screening for the function of the protein encoded by the nucleic acid and/or by hybridization with a wild-type gene.

[0050] Another embodiment of the present invention includes a recombinant nucleic acid molecule comprising a recombinant vector and a nucleic acid sequence encoding protein or peptide having at least one enzymatic activity useful for catalyzing the conversion of lignocellulosic material to fermentable sugars. According to the present invention, a recombinant vector is an engineered (i.e., artificially produced) nucleic acid molecule that is used as a tool for manipulating a nucleic acid sequence of choice and for introducing such a nucleic acid sequence into a host cell. The recombinant vector is therefore suitable for use in cloning, sequencing, and/or otherwise manipulating the nucleic acid sequence of choice, such as by expressing and/or delivering the nucleic acid sequence of choice into a host cell to form a recombinant cell. Such a vector typically contains heterologous nucleic acid sequences, that is

nucleic acid sequences that are not naturally found adjacent to nucleic acid sequence to be cloned or delivered, although the vector can also contain regulatory nucleic acid sequences (e.g., promoters, untranslated regions) which are naturally found adjacent to nucleic acid molecules of the present invention or which are useful for expression of the nucleic acid molecules of the present invention (discussed in detail below). The vector can be either RNA or DNA, either prokaryotic or eukaryotic, and typically is a plasmid. The vector can be maintained as an extrachromosomal element (e.g., a plasmid) or it can be integrated into the chromosome of a recombinant organism (e.g., a microbe or a plant). The entire vector can remain in place within a host cell, or under certain conditions, the plasmid DNA can be deleted, leaving behind the nucleic acid molecule of the present invention. The integrated nucleic acid molecule can be under chromosomal promoter control, under native or plasmid promoter control, or under a combination of several promoter controls. Single or multiple copies of the nucleic acid molecule can be integrated into the chromosome. A recombinant vector of the present invention can contain at least one selectable marker.

[0051] Typically, a recombinant nucleic acid molecule includes at least one nucleic acid molecule of the present invention operatively linked to one or more expression control sequences. According to the present invention, the phrase "operatively linked" refers to linking a nucleic acid molecule to an expression control sequence (e.g., a transcription control sequence and/or a translation control sequence) in a manner such that the molecule can be expressed when transfected (i.e., transformed, transduced, transfected, conjugated or conducted) into a host cell. Transcription control sequences are sequences that control the initiation, elongation, or termination of transcription. Particularly important transcription control sequences are those that control transcription initiation, such as promoter, enhancer, operator and repressor sequences.

[0052] Suitable transcription control sequences include any transcription control sequence that can function in a host cell or organism into which the recombinant nucleic acid molecule is to be introduced.

Enzymes and Nucleic Acids Encoding the Enzymes

[0053] As described in the examples, this invention provides several purified enzymes, including two cellobiohydrolases, (CBH Ib, SEQ ID NO. 2; CBH IIb, SEQ ID NO. 4), an endoglucanase (EG VI, SEQ ID NO. 16), a β -glucosidase (BGL, SEQ ID NO. 12), and a xylanase (Xyl II, SEQ ID NO. 18). This invention also contemplates variants of such enzymes, including variants having amino acid sequence with at least 65%, 70%, or 75% amino acid identity with these enzymes, as determined by the conventionally used BLAST algorithm.

[0054] Additionally, the invention provides the nucleic acids that encode these sequences, including gene *cbh2* (SEQ ID NO. 1, encoding CBH Ib), gene *cbh4* (SEQ ID NO. 3, encoding CBH IIb); gene *eg7* (SEQ ID NO. 15, encoding EG VI), gene *bgl1* (SEQ ID NO. 11, encoding BGL), and gene *xy12* (SEQ ID NO. 17, encoding Xyl II). This invention also contemplates variants of these nucleic acids, including variants that have at least 80%, 85% or 90% homology with these nucleic acids.

[0055] As described herein, the newly identified and isolated enzymes according to the invention can be used in conjunction with at least one other enzyme that promotes saccharification of cellulosic materials. In preferred embodiments, this additional enzyme is derived from *C. lucknowense*. For example, the enzyme may be CBH Ia (SEQ ID NO. 6), CBH IIa (SEQ ID NO. 8), EG II (SEQ ID NO. 10) or EG V (SEQ ID NO. 14). Note however, that in certain preferred embodiments, CBH Ia, CBH IIa, EG II, and EG V may be obtained by genetically modifying a microorganism or plant to express *cbh1* (SEQ ID NO. 5, encoding CBH Ia), *EG6* (SEQ ID NO. 7, encoding CBH IIa), *eg2* (SEQ ID NO. 9, encoding EG II), and/or *EG5* (SEQ ID NO. 13, encoding EG V). One particularly useful combination for saccharification is CBH Ia, CBH Ib, CBH IIb, EG II, EG V, BGL, and Xyl II.

[0056] In certain embodiments, the polynucleotides and polypeptides of the invention are evolved using molecular evolution techniques to create and to identify novel variants with desired structural, functional, and/or physical characteristics. Molecular evolution techniques can be "DNA Shuffling", or "sexual PCR" (WPC, Stemmer, PNAS, 91:10747, (1994)), also referred to as "directed molecular evolution", "exon-shuffling", "directed enzyme evolution", "in vitro evolution" and "artificial

evolution". Such reference terms are known in the art and are encompassed by the invention. Characteristics such as activity, the protein's enzyme kinetics, the protein's K_i , K_{cat} , K_m , V_{max} , K_d , thermostability, pH optimum, and the like can be modified. In certain embodiments, the polynucleotides and/or polypeptides of the invention may be evolved to confer properties that are advantageous for *in situ* enzymatic saccharification and fermentation. For example, enzymes may be evolved to perform optimally in an environment which is suitable for fermentation of sugars. In one example, the enzymes are evolved to have maximum activity in an environment with elevated temperature and high ambient alcohol content, such as an environment where an organism such as yeast is fermenting sugars. In this way, saccharification of lignocellulose and fermentation occurs in a single process step. In another example, the enzymes are evolved to resist harsh chemical or thermal environments, such as those that may be experienced during lignocellulosic pretreatments, as described herein. In these embodiments, it is not necessary to chemically or thermally pretreat the lignocellulose prior to adding enzymes. Rather, the treatment and enzymatic saccharification can be performed simultaneously. Of course, this invention also contemplates processes involving multiple steps to produce sugars from lignocellulose, such as those where evolved enzymes first saccharify lignocellulose, which is subsequently fermented by an organism, such as yeast, for example.

[0057] In other embodiments, the ability to enhance specific characteristics of a protein may also be applicable to changing the characterized activity of an enzyme to an activity completely unrelated to its initially characterized activity. Other desirable enhancements of the invention would be specific to each individual protein, and would thus be well known in the art and contemplated by the invention.

Expression of Enzymes

[0058] The microorganisms useful in the present invention and/or as a source of enzymes useful in the present invention include any microorganism producing an enzyme capable of degrading lignocellulosic material, including bacteria, yeast, and filamentous fungi. For simplicity and convenience, filamentous fungal microorganisms will be discussed herein; however, one skilled in the art will recognize that other microorganisms will be useful in the present invention. Filamentous fungi have been

widely used in industry for the production of proteins. These fungi are uniquely adapted for the production and secretion of proteins owing to their biological niche as microbial scavengers. In environments rich in biological polymers, such as forest floors, the fungi compete by secreting enzymes that degrade those polymers, producing monomers that can be readily utilized as nutrients for growth. The natural ability of fungi to produce proteins has been widely exploited, mainly for the production of industrial enzymes. Levels of protein production in natural isolates can be increased in improved strains by orders-of-magnitude; production yields of tens of grams of protein per liter of fermentation culture are commonplace.

[0059] Fungal strains, including, but not limited to, various species of *Talaromyces*, *Aspergillus*, *Trichoderma*, *Neurospora*, *Penicillium*, *Fusarium*, *Hemicola*, *Myceliophthora*, *Corynascus*, *Chaetomium*, *Tolypocladium*, *Thielavia*, *Acremonium*, *Sporotrichum*, *Thermoascus*, and *Chrysosporium*, are contemplated in the present invention. These are a few of many possible genera of fungi that will be useful sources of enzymes and/or would be suitable as host organisms for producing such enzymes mixtures. Such fungi can be obtained, for instance from various depositories such as the American Type Culture Collection (ATCC), the All Russian Collection of Microorganisms of the Russian Academy of Sciences (VKM), and Centraalbureau voor Schimmelcultures.

Mutant Strains of C. lucknowense

[0060] Particular strains of *Chrysosporium* express proteins in extremely large amounts and natural expression regulating sequences from these strains are of particular interest. These strains have been designated as *Chrysosporium* strain C1, strain UV13-6, strain NG7C-19 and strain UV18-25. They have been deposited in accordance with the Budapest Treaty with the All Russian Collection (VKM) depository institute in Moscow. The wild type C1 strain was deposited in accordance with the Budapest Treaty with the number VKM F-3500 D, deposit date August 29, 1996, C1 UV13-6 mutant was deposited with number VKM F-3632 D, and deposit date February 9, 1998, C1 NG7c-19 mutant was deposited with number VKM F-3633 D and deposit date

February 9, 1998 and C1 UV18-25 mutant was deposited with number VKM F-3631 D and deposit date February 9, 1998.

[0061] Preferably an expression-regulating region enabling high expression in the selected host is applied. This can also be a high expression-regulating region derived from a heterologous host, such as are well known in the art. Specific examples of proteins known to be expressed in large quantities and thus providing suitable expression regulating sequences for the invention are without being limited thereto hydrophobin, protease, amylase, xylanase, pectinase, esterase, beta-galactosidase, cellulase (e.g. endo-glucanase, cellobiohydrolase) and polygalacturonase. The high production has been ascertained in both solid state and submerged fermentation conditions. Assays for assessing the presence or production of such proteins are well known in the art.

[0062] Heterologous expression-regulating sequences also work efficiently in *Chrysosporium* as native *Chrysosporium* sequences. This allows well known constructs and vectors to be used in transformation of *Chrysosporium* as well as offering numerous other possibilities for constructing vectors enabling good rates of expression in this novel expression and secretion host. As extremely high expression rates for cellulase have been ascertained for *Chrysosporium* strains, the expression regulating regions of such proteins are particularly preferred.

[0063] A nucleic acid construct comprising a nucleic acid expression regulatory region from *Chrysosporium lucknowense* or a derivative thereof forms a separate embodiment of the invention as does the mutant *Chrysosporium* strain comprising such regions operably linked to a gene encoding a polypeptide to be expressed. In preferred embodiments, such a nucleic acid construct will be an expression regulatory region from *Chrysosporium* associated with cellobiohydrolase, endoglucanase, β -glucosidase, and/or xylanase expression.

[0064] The invention also covers genetically engineered *Chrysosporium* strains wherein the sequence that is introduced can be of *Chrysosporium* origin. Such a strain can, however, be distinguished from natively occurring strains by virtue of for example heterologous sequences being present in the nucleic acid sequence used to transform or transfect the *Chrysosporium*, by virtue of the fact that multiple copies of the sequence

encoding the polypeptide of interest are present or by virtue of the fact that these are expressed in an amount exceeding that of the non-engineered strain under identical conditions or by virtue of the fact that expression occurs under normally non-expressing conditions. The latter can be the case if an inducible promoter regulates the sequence of interest contrary to the non-recombinant situation or if another factor induces the expression than is the case in the non-engineered strain. The invention as defined in the preceding embodiments is not intended to cover naturally occurring *Chrysosporium* strains. The invention is directed at strains derived through engineering either using classical genetic technologies or genetic engineering methodologies.

[0065] A method of production of a recombinant microorganism or plant is also part of the subject invention. The method comprises stably introducing a nucleic acid sequence encoding a heterologous or homologous polypeptide into a microbial strain or plant, the nucleic acid sequence being operably linked to an expression regulating region. Such procedures for transforming filamentous fungi have been previously reported. In one preferred embodiment, the mutant *Chrysosporium lucknowense* is derived from UV18-25 (Acc. No. VKM F-3631 D) that has been engineered to overexpress the Xyl II gene.

Genetically Modified Organisms

[0066] As used herein, a genetically modified microorganism can include a genetically modified bacterium, yeast, fungus, or other microbe. Such a genetically modified microorganism has a genome which is modified (i.e., mutated or changed) from its normal (i.e., wild-type or naturally occurring) form such that a desired result is achieved (e.g., increased or modified activity and/or production of a least one enzyme or a multi-enzyme product for conversion of lignocellulosic material to fermentable sugars). Genetic modification of a microorganism can be accomplished by using classical strain development and/or molecular genetic techniques. Such techniques known in the art and are generally disclosed for microorganisms, for example, in Sambrook et al., 1989, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Labs Press. The reference Sambrook et al., *ibid.*, is incorporated by reference herein in its entirety. A genetically modified microorganism can include a microorganism in

which nucleic acid molecules have been inserted, deleted or modified (i.e., mutated; e.g., by insertion, deletion, substitution, and/or inversion of nucleotides), in such a manner that such modifications provide the desired effect within the microorganism.

[0067] In one aspect of the invention, a genetically modified microorganism can endogenously contain and express an enzyme or a multi-enzyme product for the conversion of lignocellulosic material to fermentable sugars, and the genetic modification can be a genetic modification of one or more of such endogenous enzymes, whereby the modification has some effect on the ability of the microorganism to convert lignocellulosic material to fermentable sugars.

[0068] In another aspect of the invention, a genetically modified microorganism can endogenously contain and express an enzyme or a multi-enzyme product for the conversion of lignocellulosic material to fermentable sugars, and the genetic modification can be an introduction of at least one exogenous nucleic acid sequence (e.g., a recombinant nucleic acid molecule), wherein the exogenous nucleic acid sequence encodes at least one additional enzyme useful for the conversion of lignocellulosic material to fermentable sugars and/or a protein that improves the efficiency of the enzyme or multi-enzyme product for the conversion of lignocellulosic material to fermentable sugars. In this aspect of the invention, the microorganism can also have at least one modification to a gene or genes comprising its endogenous enzyme(s) for the conversion of lignocellulosic material to fermentable sugars.

[0069] In yet another aspect of the invention, the genetically modified microorganism does not necessarily endogenously (naturally) contain an enzyme or a multi-enzyme product for the conversion of lignocellulosic material to fermentable sugars, but is genetically modified to introduce at least one recombinant nucleic acid molecule encoding at least one enzyme, a multiplicity of enzymes, or a multi-enzyme product for the conversion of lignocellulosic material to fermentable sugars. Such a microorganism can be used in a method of the invention, or as a production microorganism for crude fermentation products, partially purified recombinant enzymes, and/or purified recombinant enzymes, any of which can then be used in a method of the present invention.

Genetically Modified Plants

[0070] The invention also contemplates genetically modified plants comprising such genes. The plants may be used for production of the enzymes, or as the lignocellulosic material used as a substrate in the methods of the invention. Methods to generate recombinant plants are known in the art. For instance, numerous methods for plant transformation have been developed, including biological and physical transformation protocols. See, for example, Miki et al., "Procedures for Introducing Foreign DNA into Plants" in *Methods in Plant Molecular Biology and Biotechnology*, Glick, B.R. and Thompson, J.E. Eds. (CRC Press, Inc., Boca Raton, 1993) pp. 67 -88. In addition, vectors and in vitro culture methods for plant cell or tissue transformation and regeneration of plants are available. See, for example, Gruber et al., "Vectors for Plant Transformation" in *Methods in Plant Molecular Biology and Biotechnology*, Glick, B.R. and Thompson, J.E. Eds. (CRC Press, Inc., Boca Raton, 1993) pp. 89-119.

[0071] In certain embodiments of the invention, genetically modified plants that express the enzymes of this invention are obtained by introducing an expression vector into plants based on the natural transformation system of *Agrobacterium*. See, for example, Horsch et al., *Science*, 227:1229 (1985). *A. tumefaciens* and *A. rhizogenes* are plant pathogenic soil bacteria which genetically transform plant cells. The Ti and Ri plasmids of *A. tumefaciens* and *A. rhizogenes*, respectively, carry genes responsible for genetic transformation of the plant. See, for example, Kado, C.I., *Crit. Rev. Plant. Sci.* 10:1 (1991). Descriptions of *Agrobacterium* vector systems and methods for *Agrobacterium*-mediated gene transfer are provided by numerous references, including Gruber et al., *supra*, Miki et al., *supra*, Moloney et al., *Plant Cell Reports* 8:238 (1989), and U.S. Patents Nos. 4,940,838 and 5,464,763, hereby incorporated by reference in their entirety.

[0072] In other embodiments, genetically modified plants are obtained by microprojectile-mediated transformation wherein DNA is carried on the surface of microprojectiles. The expression vector is introduced into plant tissues with a biolistic device that accelerates the microprojectiles to speeds sufficient to penetrate plant cell walls and membranes. Sanford et al., *Part. Sci. Technol.* 5:27 (1987), Sanford, J.C.,

Trends Biotech. 6:299 (1988), Sanford, J.C., *Physiol. Plant* 79:206 (1990), Klein et al., *Biotechnology* 10:268 (1992).

[0073] Another method for physical delivery of DNA to plants contemplated by this invention is sonication of target cells. Zhang et al., *Bio Technology* 9:996 (1991). Alternatively, liposome or spheroplast fusion have been used to introduce expression vectors into plants. Deshayes et al., *EMBO J.*, 4:2731 (1985), Christou et al., *Proc Natl. Acad. Sci. USA* 84:3962 (1987). Direct uptake of DNA into protoplasts using CaCl₂ precipitation, polyvinyl alcohol or poly-L-ornithine have also been reported. Hain et al., *Mol. Gen. Genet.* 199:161 (1985) and Draper et al., *Plant Cell Physiol.* 23:451 (1982). Electroporation of protoplasts and whole cells and tissues have also been described. Donn et al., In *Abstracts of VIIIth International Congress on Plant Cell and Tissue Culture* IAPTC, A2-38, p. 53 (1990); D'Halluin et al., *Plant Cell* 4:1495-1505 (1992) and Spencer et al., *Plant Mol. Biol.* 24:51-61 (1994).

Methods of Using the Enzymes and Mutant Strains of C. lucknowense

[0074] This invention also provides methods of enzymatic saccharification of cellulosic materials. Any cellulose containing material can be treated by the enzymes of this invention, non-limiting examples of which include orchard prunings, chaparral, mill waste, urban wood waste, yard waste, municipal waste, logging waste, forest thinnings, short-rotation woody crops, industrial waste, wheat straw, oat straw, rice straw, barley straw, rye straw, flax straw, sugar cane, corn stover, corn stalks, corn cobs, corn husks, prairie grass, gamagrass, foxtail; sugar beet pulp, citrus fruit pulp, seed hulls, cellulosic animal wastes, lawn clippings, cotton, and seaweed.

[0075] In certain preferred embodiments, the lignocellulosic materials are pretreated before being exposed to the enzymes or enzyme mixtures of the invention. Generally speaking, the pretreatment can be any procedure that makes the subsequent enzymatic saccharification of the lignocellulosic materials more efficient (i.e., either less time-consuming or less costly). For example, the lignocellulosic material may be pretreated by methods including, but not limited to, exposure to acids, bases, solvents, heat, peroxides, ozone, or some combination thereof prior to enzymatic saccharification. These pretreatments can also be combined with other forms of

processing, such as mechanical shredding, grinding, milling, or rapid depressurization (e.g. steam explosion).

[0076] Generally, enzymatic saccharification according to the invention involves using CBH Ia, CBH Ib, EG VI, BGL, Xyl II, or mixtures thereof. One or more of these enzymes may be further combined with other enzymes capable of promoting enzymatic saccharification, which may be derived from *C. lucknowense*, a mutant strain, or another organism. For example, in one embodiment, the enzymatic saccharification involves an enzyme mixture comprising CBH Ia, CBH Ib, CBH IIb, EG II, EG V, BGL, and Xyl II. In other preferred embodiments, the enzymatic mixture contains a cellobiohydrolase, which may be CBH Ia, CBH Ib, CBH IIa, CBH IIb, and mixtures thereof, with a β -glucosidase such as BGL.

[0077] In certain embodiments, the enzyme compositions are artificial enzyme compositions that contain purified forms of CBH Ia, CBH Ib, CBH IIb, EG II, EG VI, BGL, or Xyl II. The purified forms of these enzymes may be used alone or mixed together. In certain preferred embodiments, the selected purified enzymes are present in higher relative amounts than would be the case for the enzyme secretions of the wild type *C. lucknowense*.

[0078] In certain embodiments, the invention provides a mutant strain of *C. lucknowense* that is capable of expressing CBH Ia, CBH Ib, CBH IIa, CBH IIb, EG II, EG V, EG VI, BGL, or Xyl II, or mixtures thereof in proportions higher than found in the enzyme secretions of the wild-type organism. The secreted enzymes of such a mutant strain of *C. lucknowense* may serve as a raw source from which purified forms of CBH Ia, CBH Ib, CBH IIa, CBH IIb, EG II, EG V, EG VI, BGL, or Xyl II, can be produced. Alternatively, the secreted enzymes of such a mutant strain may also be applied directly to the cellulosic materials to be saccharified. In particularly preferred embodiments, the cellulosic materials are exposed directly to the mutant strain of *C. lucknowense* in an environment conducive to the proliferation of the mutant strain of *C. lucknowense*, such as in a bioreactor. The in situ secretions of CBH Ia, CBH Ib, CBH IIa, CBH IIb, EG II, EG V, EG VI, BGL, or Xyl II, or mixtures thereof by the mutant strain of *C. lucknowense*, in proportions higher than found in the enzyme secretions of the wild-type organism, lead to enhanced in situ saccharification of the cellulosic material.

[0079] Following enzymatic treatment by the inventive enzymatic compositions of the invention, the fermentable sugar that is produced can be exposed to microorganisms, either naturally occurring or genetically engineered, that are capable of fermenting the sugar to produce ethanol or some other value-added fermentation product. Preferably, substantially all of the glucose is converted to ethanol, which may be subsequently used as a fuel, solvent, or chemical reactant. In preferred embodiments, the ethanol is used as a fuel for powering transportation vehicles, non-limiting examples of which include cars, trucks, buses, mopeds and motorcycles. Other potential fermentation products from glucose include, but are not limited to, biofuels (including ethanol); lactic acid; plastics; specialty chemicals; organic acids, including citric acid, succinic acid and maleic acid; solvents; animal feed supplements; pharmaceuticals; vitamins; amino acids, such as lysine, methionine, tryptophan, threonine, and aspartic acid; industrial enzymes, such as proteases, cellulases, amylases, glucanases, lactases, lipases, lyases, oxidoreductases, and transferases; and chemical feedstocks.

EXAMPLES

Example 1: Enzyme Isolation

[0080] Culture filtrates produced by the *C. lucknowense* mutant strains were used for isolation of individual enzymes. Commercial preparation of NCE-L600 (*C. lucknowense*) were from Dyadic International, Inc., USA.

[0081] Highly purified BGL (cellobiase) from *Aspergillus japonicus* was obtained from a commercial preparation, having specific cellobiase activity 50 U mg⁻¹ protein (pH 5.0, 40°C), and was used in the experiments on hydrolysis of insoluble cellulose.

Example 2: Enzyme purification

[0082] The enzyme purification was carried out by chromatography on a Pharmacia FPLC system (Sweden). Cellobiohydrolases and endoglucanases BGL and Xyl II were isolated from a *C. lucknowense* UV18-25 culture filtrate. BGL and Xyl II

(xylanase II) were isolated from culture filtrates produced by the *C. lucknowense* UV18ΔCbh1#10 and Xyl2-18 mutant strains, respectively.

[0083] In all cases, the first purification stage was anion-exchange chromatography on a Source 15Q column (40 ml volume). The column was equilibrated with 0.02 M Bis-Tris-HCl buffer, pH 6.8. The initial culture filtrate was preliminarily desalted and transferred into the starting buffer by gel-filtration on Acrylex P4 (Reanal, Hungary). The sample (400 mg of protein) was applied to the Source 15Q column, and the elution was carried out with a gradient of 0-1 M NaCl at a flow rate of 10 ml min⁻¹.

[0084] The first protein fraction after the Source 15Q, eluted at 0.05 M NaCl and having high Avicelase activity, was subjected to hydrophobic interaction chromatography on a Source 15 Isopropyl column (Pharmacia, Sweden). The column was equilibrated with 1.7 M ammonium sulfate in 50 mM Na-acetate buffer, pH 5.0. Proteins were eluted with a reverse linear gradient of 1.7-0 M ammonium sulfate at a flow rate of 4 ml min⁻¹. The protein fraction with the highest activity against Avicel (eluting at a salt concentration of 0.30-0.35 M) contained the homogeneous protein with a molecular mass of 70 kDa (CBH IIb, see Fig. 1).

[0085] The protein fraction after the Source 15Q, eluted at 0.22 M NaCl and having the activity against Avicel and *p*-NP-β-D-cellobioside, was further purified by chromatofocusing on a Mono P HR 5/20 column (Pharmacia, Sweden). The column was equilibrated with 0.025 M Na-formate buffer, pH 4.0. Proteins were eluted with a gradient of pH 4.5-3.0 (using Polybuffer 74) at a flow rate of 0.5 ml min⁻¹. Homogeneous 60 kDa CBH Ib was obtained as a result of chromatofocusing (FIG. 1).

[0086] The two newly isolated cellobiohydrolases are homogeneous according to the data of SDS-PAGE and isoelectrofocusing (FIG. 1), their molecular masses were found to be 60 and 70 kDa, pI 3.8 and 5.6, respectively. Peptide mass fingerprinting using MALDI-TOF mass spectrometry (data not shown) indicated that these proteins were different from the above-mentioned cellobiohydrolases (Cel6A and Cel7A) as well as from other *C. lucknowense* enzymes previously isolated. Subsequent *de novo* sequencing of tryptic peptides from the new cellobiohydrolases, using tandem TOF/TOF mass spectrometry (MS/MS), followed by the BLAST search in the SWISS-PROT (UniProtKB) database showed that the 60 kDa and 70 kDa proteins display

sequence similarity to cellobiohydrolases from the GH families 7 and 6 (Table 1, see classification into families in [HyperTextTransferProtocol://afmb.cnrs-mrs.fr/CAZY/](http://afmb.cnrs-mrs.fr/CAZY/)). So, they were classified as Cel7B (CBH Ib) and Cel6B (CBH IIb), respectively. Thus, the *C. lucknowense* fungus secretes at least four cellobiohydrolases encoded by different genes, two of them belonging to the glycosyl hydrolase family 6 (GH6) and two other enzymes – to the GH7 family (Table 2). The molecules of the CBH Ia (Cel7A) and CBH IIb (Cel6B) represent typical cellulases consisting of a catalytic domain and CBM connected by a flexible peptide linker. The molecules of CBH Ib (Cel7B) and CBH IIa (Cel6A) consist of only the catalytic domains (they lack CBM). It should be noted that the most studied fungus *T. reesei* has only two cellobiohydrolases: I (Cel7A) and II (Cel6A). Other fungi, such as *Humicola insolens*, also secrete two cellobiohydrolases (Cel7A and Cel6A), while *Phanerochaete chrysosporium* produces at least seven different cellobiohydrolases, of which six enzymes belong to the GH7 family. All the enzymes mentioned, except for the *P. chrysosporium* CBH 1-1 (Cel7A), possess CBM.

[0087] The BGL was isolated from the protein fraction after the Source 15Q (eluted at 0.10 M NaCl) containing the highest activity against *p*-NP- β -D-glucopyranoside and cellobiose. The fraction was subjected to hydrophobic interaction chromatography as described above, the homogeneous BGL with a molecular mass of 106 kDa and *pI* 4.8 was eluted at 1.3 M of ammonium sulfate. The specific activity of the BGL toward *p*-NP- β -D-glucopyranoside and cellobiose was found to be 11 and 26 U mg⁻¹ of protein, respectively (40°C, pH 5.0). Purified BGL had optimum activity at pH 4.0 and retained >50% of activity in the range of pH 2.5-6.5. The temperature optimum was 40°C. After heating for three hours, the enzyme retained 10% activity at 60°C, 64% at 50°C, and 100% at 40°C. The enzyme was highly active against cellobiose, gentiobiose, and laminarobiose as substrates. Weak activity was also observed using sophorose, cellotriose, cellotetraose, cellopentaose, and cellohexaose as substrates. No activity was observed with lactose or trehalose as substrates.

[0088] The homogeneous Xyl II (24 kDa, *pI* 7.9) was obtained after anion-exchange chromatography followed by hydrophobic interaction chromatography as described above and gel-filtration on a Superose 12 HR 10/30 column (Pharmacia, Sweden). Elution at the last chromatographic stage was performed with 0.1 M Na-acetate buffer, pH 5.0, at a flow rate of 0.3 ml min⁻¹. The Xyl II had specific xylanase

activity of 395 U mg⁻¹ of protein (50°C, pH 5.0, birchwood xylan as a substrate). The enzyme had a pH optimum of 6.0 and a temperature optimum of 70°C. Xyl II was highly specific for xylan as substrate, with no activity against carboxymethylcellulose (CMC) or barley β -glucan.

[0089] The *C. lucknowense* CBH Ia (65 kDa), CBH IIa (43 kDa), EG II (51 kDa), EG V (25 kDa), EG VI (47 kDa) were purified as described elsewhere (see, Gusakov AV, Sinitsyn AP, Salanovich TN, Bukhtojarov FE, Markov AV, Ustinov BB, van Zeijl C, Punt P, Burlingame R. "Purification, cloning and characterisation of two forms of thermostable and highly active cellobiohydrolase I (Cel7A) produced by the industrial strain of *Chrysosporium lucknowense*" *Enzyme Microb Technol* 2005;36:57-69; Bukhtojarov FE, Ustinov BB, Salanovich TN, Antonov AI, Gusakov AV, Okunev ON, Sinitsyn AP. "Cellulase complex of the fungus *Chrysosporium lucknowense*: isolation and characterization of endoglucanases and cellobiohydrolases", *Biochemistry* (Moscow) 2004;69:542-51.

[0090] The enzyme purity was characterized by SDS-PAGE and isoelectrofocusing. SDS-PAGE was carried out in 12% gel using a Mini Protean II equipment (Bio-Rad Laboratories, USA). Isoelectrofocusing was performed on a Model 111 Mini IEF Cell (Bio-Rad Laboratories, USA). Staining of protein was carried out with Coomassie Blue.

Example 3: MALDI-TOF and tandem TOF/TOF mass spectrometry of peptides

[0091] The in-gel tryptic digestion of the protein bands after the SDS-PAGE was carried out essentially as described by Smith (Smith BE. Protein sequencing protocols. Totowa: Humana Press; 1997). Trypsin (Promega, modified, 5 μ g/mL) in 50 mM NH₄HCO₃ was used for a protein digestion. The resulting peptides were extracted from a gel with 20% aqueous acetonitrile containing 0.1% trifluoroacetic acid and subjected to MALDI-TOF MS (see, James P. (Ed.) Proteome research: mass spectrometry. Berlin: Springer-Verlag; 2001.) Selected peptides from the mass spectra of the tryptic digests of the CBH Ib and IIb were analyzed by tandem mass spectrometry in order to determine their sequences *de novo*. Ultraflex TOF/TOF mass spectrometer (Bruker Daltonik GmbH, Germany) was used in the MS experiments.

Example 4: Enzyme activity assays

[0092] CMCase activity was measured by assaying reducing sugars released after 5 min of enzyme reaction with 0.5% carboxymethylcellulose (CMC, medium viscosity, Sigma, USA) at pH 5.0 and 50°C (Sinitsyn AP, Chernoglazov VM, Gusakov AV. "Methods of investigation and properties of cellulolytic enzymes" (in Russian), *Biotechnology Series*, v. 25. Moscow: VINITI Press; 1990). Enzyme activities against barley β -glucan (Megazyme, Australia) and birchwood xylan (Sigma, USA) were determined in the same way as the CMCase activity, except the incubation time was 10 min. Avicelase activity was determined by analysing reducing sugars released after 60 min of enzyme reaction with 5 mg ml⁻¹ Avicel PH 105 (Serva, Germany) at pH 5.0 and 40°C. Reducing sugars were analysed by the Somogyi-Nelson method (Sinitsyn AP, Chernoglazov VM, Gusakov AV, "Methods of investigation and properties of cellulolytic enzymes" (in Russian); *Biotechnology Series*, v. 25. Moscow: VINITI Press; 1990; Somogyi M., "Notes on sugar determination" *J Biol Chem* 1952;195:19-23.. Filter paper activity (FPA) was determined as recommended by Ghose (Ghose TK. "Measurement of cellulase activities", *Pure Appl Chem* 1987;59:257-68).

[0093] Activities against *p*-NP- β -D-glucopyranoside, *p*-NP- β -D-cellobioside and *p*-NP- β -D-lactoside (Sigma, USA) were determined at pH 5.0 and 40°C as described elsewhere (Gusakov AV, Sinitsyn AP, Salanovich TN, Bukhtojarov FE, Markov AV, Ustinov BB, van Zeijl C, Punt P, Burlingame R. "Purification, cloning and characterisation of two forms of thermostable and highly active cellobiohydrolase I (Cel7A) produced by the industrial strain of *Chrysosporium lucknowense*", *Enzyme Microb Technol* 2005;36:57-69).

[0094] . Cellobiase activity was assayed at pH 5.0 and 40°C by measuring the initial rate of glucose release from 2 mM cellobiose by the glucose oxidase – peroxidase method (Sinitsyn AP, Chernoglazov VM, Gusakov AV, "Methods of investigation and properties of cellulolytic enzymes" (in Russian), *Biotechnology Series*, v. 25. Moscow: VINITI Press; 1990)..

[0095] All activities were expressed in International Units, i. e. one unit of activity corresponded to the quantity of enzyme hydrolysing one μ mol of substrate or releasing one μ mol of reducing sugars (in glucose equivalents) per one minute.

Example 5: Enzymatic hydrolysis of cellulosic substrates

[0096] The enzymatic hydrolysis of cellulosic substrates was carried out at pH 5.0 under magnetic stirring. Avicel PH 105 (Serva, Germany), cotton pretreated with acetone-ethanol mixture (1:1) for two days in order to remove wax from the surface of cellulose fibres, and Douglas fir wood pretreated by organosolv were used as substrates.

[0097] The experiments on progress kinetics of Avicel hydrolysis by purified individual cellobiohydrolases and experiments on synergistic interaction between *C. lucknowense* cellulases (with cotton as a substrate) were carried out at 40°C. The substrate concentration in those experiments was 5 mg ml⁻¹. In order to eliminate the effect of product (cellobiose) inhibition on the kinetics and to convert all cellooligosaccharides to glucose, the hydrolysis was carried out in the presence of purified BGL (cellobiase) from *A. japonicus*, which was extra added to the reaction system in excessive quantity (0.5 U ml⁻¹).

[0098] The experiments on enzymatic saccharification of Avicel, cotton, and pretreated Douglas fir wood by combinations of purified *C. lucknowense* enzymes and crude multienzyme preparations were carried out at 50°C. The concentration of Avicel and pretreated wood in those experiments was 50 mg ml⁻¹, while the concentration of cotton was 25 mg ml⁻¹.

[0099] A typical experiment was carried out in the following way. A weighed amount of dry cellulosic substrate was placed into a 2-ml plastic test tube, then 0.5-1 ml of 0.05 M Na-acetate buffer, containing 1 mM NaN₃ to prevent microbial contamination, was added, and the substrate was soaked in the buffer for 1 h. Then, the tube was placed into a thermostated water bath, located on a magnetic stirrer, and suitably diluted enzyme solution in the same buffer was added to the substrate suspension in order to adjust the total volume of the reaction system to 2 ml and to start the hydrolysis. The tube was hermetically closed with a lid, and the hydrolysis was carried out with magnetic stirring. At defined times in the reaction, an aliquot of the suspension (0.05-0.1 ml) was taken, diluted, centrifuged for 3 min at 15000 rpm, and the concentrations of glucose and reducing sugars in the supernatant were determined by the glucose oxidase – peroxidase and Somogyi-Nelson methods. In those cases,

when glucose was a single product of the reaction, the degree of substrate conversion (for Avicel and cotton, which represented pure cellulosic substrates) was calculated using the following equation:

$$\text{Conversion (\%)} = \frac{\text{Glucose concentration (mg ml}^{-1}\text{)} \times 100\%}{\text{Initial substrate concentration (mg ml}^{-1}\text{)} \times 1.11}$$

[00100] The kinetic experiments were carried out in duplicates. Protein concentration was the measure of enzyme loading in the reaction system. In the case of purified enzymes, the protein concentration was calculated from the UV absorption at 280 nm using enzyme extinction coefficients predicted by the ProtParam tool (<http://WorldWideWeb.expasy.ch/tools/protparam.HyperTextMarkupLanguage>). For crude multienzyme preparations, the protein concentration was determined by the Lowry method using bovine serum albumin as a standard.

[00101] The CBH Ib and IIb displayed maximum activity at pH 4.7 and 5.0. Both enzymes were stable during 24 h incubation at pH 5.0 and 50°C. Study of the enzyme adsorption on Avicel, carried out at pH 5.0 and 6°C, revealed that only the CBH IIb has CBM. After incubation of the CBH Ib and IIb (1 mg ml⁻¹) with Avicel (25 mg ml⁻¹) for 30 min on stirring the degree of protein adsorption was 65 and 99%, respectively. It should be noted that the adsorption degree of the catalytic domain of the *C. lucknowense* CBH Ia was 59% under the same conditions, while that for the full size *C. lucknowense* CBH Ia (an enzyme with CBM) was 89%.

[00102] The CBH IIb had a high activity against Avicel and very low CMCase activity, while the activity toward synthetic *p*-nitrophenyl derivatives of disaccharides was completely absent (Table 2). The CBH Ib displayed lower Avicelase activity, but hydrolysed *p*-NP-β-D-cellobioside and *p*-NP-β-D-lactoside, which is typical for family 7 cellulases. For a comparison, specific activities of previously isolated *C. lucknowense* cellobiohydrolases (now named as CBH Ia and CBH IIa) are also given in Table 2.

[00103] FIG. 2 shows the progress kinetics of Avicel hydrolysis by the all purified *C. lucknowense* cellobiohydrolases, where the enzymes were equalized by protein concentration (0.1 mg ml⁻¹). In order to eliminate the effect of product

(cellobiose) inhibition on the kinetics, the hydrolysis was carried out in the presence of purified BGL (cellobiase) from *A. japonicus*, added to the reaction system in excessive quantity (0.5 U ml^{-1}).

[00104] The highest hydrolysis rate amongst a few cellobiohydrolases tested, including three other *C. lucknowense* enzymes (CBH Ia, Ib, IIa) was observed in the case of *C. lucknowense* CBH IIb: 3.2 mg ml^{-1} of glucose, i.e. 58% cellulose conversion was achieved after 5 days of hydrolysis (see **FIG. 2**). The *C. lucknowense* CBH Ia (which has a CBM) was notably less effective (the yield of glucose after 5 days was 2.5 mg ml^{-1} , which corresponded to the cellulose conversion degree of 46%, respectively). As expected, the *C. lucknowense* cellobiohydrolases without CBM (CBH Ib and IIa) had the lowest ability to hydrolyse Avicel: only 23 and 21% cellulose conversion was achieved after the same time of reaction.

[00105] Both *C. lucknowense* cellobiohydrolases having a CBM (Ia and IIb) displayed a pronounced synergism with three major endoglucanases from the same fungus (EG II, EG V, EG VI) in hydrolysis of cotton as well as a strong synergy with each other (Table 3). In these studies, the concentration of cotton was 5 mg ml^{-1} , the CBH concentration was 0.15 mg ml^{-1} in all cases, while the EG concentration was always 0.05 mg ml^{-1} . In order to eliminate the effect of product inhibition on the kinetics and to convert the intermediate oligosaccharides to glucose, the hydrolysis was carried out in the presence of purified BGL from *A. japonicus*, added to the reaction system in excessive quantity (0.5 U ml^{-1}). The experiments were carried out at pH 5.0 and 40°C for 140 h.

[00106] As seen from Table 3, individual cellobiohydrolases, CBH Ia and CBH IIb, and the individual endoglucanases, did not completely hydrolyze cotton under the conditions tested. The CBH IIb provided the highest glucose yield after 140 h of hydrolysis: 1.18 mg ml^{-1} , which corresponded to the substrate conversion degree of 21%. However, when either cellobiohydrolase was incubated with endoglucanase, a pronounced synergism was observed. The highest glucose yields ($4.1\text{--}4.7 \text{ mg ml}^{-1}$) were achieved with combinations of CBH Ia or CBH IIb with EG II, the coefficient of synergism being varied in the range of 2.6–2.8. A strong synergism ($K_{\text{syn}} = 2.75$) was also observed between CBH Ia and CBH IIb. In fact, the combination of two

cellobiohydrolases (1:1 by weight) with BGL provided practically complete conversion (98.6%) of cotton cellulose to glucose after 140 h of hydrolysis.

[00107] As an example, the progress kinetics of cotton hydrolysis by combinations of CBH IIb with other *C. lucknowense* enzymes are shown in FIG. 3, where real experimental data are shown with open symbols (continuous curves) while the theoretical sums of glucose concentrations obtained under the action of individual enzymes are shown with filled symbols (dotted lines). Glucose yields obtained after 140 h of cotton hydrolysis under the action of individual cellobiohydrolases and endoglucanases and their combinations are summarized in Table 3. The coefficient of synergism (K_{syn}) was calculated as a ratio of experimental glucose concentration (column 2 of Table 3) to the theoretical sum of glucose concentrations (column 3).

[00108] Using four purified *C. lucknowense* enzymes (CBH Ia and IIb, EG II, BGL), an artificial cellulase complex was constructed (*C.l.* combination #1) that demonstrated an extremely high ability to convert different cellulosic substrates to glucose (FIGS. 4-6). This multienzyme composition was notably more effective in hydrolysis of pure crystalline cellulose (cotton and Avicel) than the crude *C. lucknowense* multienzyme preparation NCE-L600. In 72-h hydrolysis of a lignocellulosic substrate (Douglas fir wood pretreated by organosolv), the *C.l.* combination #1 was also very effective in cellulose hydrolysis.

[00109] In *C. lucknowense* combination #1, the enzyme consisted of the two cellobiohydrolases CBH Ia and CBH Ib, and the endoglucanase EG II, the enzymes with strong adsorption ability on crystalline cellulose (the molecules of these enzymes have CBM). The activity of tightly adsorbed cellulases is gradually decreased during in the course of hydrolysis of insoluble cellulose as a result of the enzyme limited mobility along the substrate surface or unproductive binding (so called pseudoinactivation). Without wishing to be bound by theory, it is believed that there may exist a synergism between tightly and loosely adsorbed cellulases wherein loosely binding cellulases (enzymes without CBM) may destroy obstacles hindering the processive action of the tightly adsorbed cellobiohydrolases, thus helping them to move to the next cellulose reactive sites. The total protein concentration in the reaction system was 0.5 mg ml⁻¹. The composition of the multienzyme composition (*C.l.* combination #1) was the

following: 0.2 mg ml⁻¹ of CBH Ia + 0.2 mg ml⁻¹ of CBH IIb + 0.08 mg ml⁻¹ of EG II + 0.02 mg ml⁻¹ of BGL. Avicel (50 mg ml⁻¹) and cotton (25 mg ml⁻¹) were used as substrates representing pure crystalline cellulose in these experiments. Sample of Douglas fir wood pretreated by organosolv (50 mg ml⁻¹) was taken as an example of real lignocellulosic feedstock that may be used for bioconversion to ethanol. A crude *C. lucknowense* multienzyme cellulase preparation NCE L-600 (diluted so that the protein concentration in the reaction system would also be 0.5 mg ml⁻¹) was taken for a comparison in these studies. The hydrolysis experiments with them were carried out also in the presence of extra added *A. japonicus* BGL (0.5 U ml⁻¹).

[00110] The progress kinetics of cotton, Avicel and Douglas fir hydrolysis by different cellulase multienzyme preparations are shown in FIGS. 4-6. It should be noted that in all cases, the concentrations of glucose and reducing sugars after 24-72 h of hydrolysis in a concrete experiment were practically the same, i.e. glucose made up >96% of the total soluble sugars. So, the glucose yield can be taken as reliable criterion in comparison of the hydrolytic efficiency of different multienzyme samples.

[00111] In hydrolysis of cotton (FIG. 4), the combination #1 of purified *C. lucknowense* enzymes provided much higher glucose yield after 72 h of the reaction (23.4 mg ml⁻¹, i.e. 84% degree of substrate conversion) than the 4.2 mg ml⁻¹ exhibited by (NCE-L600). In hydrolysis of Avicel (FIG. 5), the *C.l.* combination #1 was also superior (45.0 mg ml⁻¹ of glucose, or 81% substrate conversion after 72 h of hydrolysis). In the case of pretreated Douglas fir (FIG. 6), the *C.l.* combination #1 was also effective (28.8 mg ml⁻¹ glucose, 63% conversion after 72 hours).

[00112] Unlike Avicel and cotton, the pretreated wood sample contained not only cellulose (~85%) but also lignin (13%) and hemicellulose (2%). The artificial *C. lucknowense* four-enzyme combination #1 was composed of only cellulases; all of them, except for the BGL, having CBM. All other multienzyme samples possessed not only cellulase but also xylanase and other types of carbohydrase activity, i.e. they contained non-cellulase accessory enzymes. This may explain relatively lower efficiency of the *C.l.* combination #1 on pretreated Douglas fir compared to the *P. verruculosum* #151 preparation (FIG. 6).

[00113] In one set of experiments (FIG. 7), the pretreated wood sample was hydrolysed by different compositions of purified *C. lucknowense* enzymes, to which cellulases lacking a CBM were included (EG V or EG V in combination with CBH Ib). The total protein concentration in the reaction system was maintained at the same level of 0.5 mg ml⁻¹ (Table 5). Indeed, two *C.l.* combinations (#3 and #4), containing weakly adsorbed enzymes, provided a notable enhancement of the glucose yield after 72 h of the enzymatic reaction in comparison with the *C.l.* combination #1.

[00114] In two experiments, the highly active *C. lucknowense* Xyl II (Xyn11A) was added to the above-mentioned four enzymes (*C.l.* combinations #2 and #4). Since a synergism between tightly and loosely adsorbed cellulases has been described [38], EG V or EG V together with CBH Ib (both enzymes have lack CBM) were used in the *C.l.* combinations #3 and #4.

[00115] As can be seen from FIG. 7, the initial rate of glucose formation decreased sequentially from *C.l.* combination #1 to combination #4, however the glucose yield after 2-3 days of hydrolysis increased in the same sequence. The Xyl II demonstrated only slight positive effect on the glucose yield, while the EG V or EG V together with CBH Ib provided a very notable increase in the product concentration after 72 h hydrolysis of wood (37 and 41 mg ml⁻¹, respectively) compared to the *C.l.* combination #1 (29 mg ml⁻¹), i.e. the combinations #3 and #4 performed much better than all crude multienzyme samples (FIG. 6).

[00116] The low performance of the crude *C. lucknowense* preparation (NCE-L600) in hydrolysis of different cellulosic substrates (FIGS. 4-6) deserves a special attention. Without wishing to be bound by theory, it may be explained by the low total content of different cellobiohydrolases in the NCE-L600 (35-40% of the total protein content). Moreover, two of four *C. lucknowense* cellobiohydrolases (Ib and IIa) lack CBM, while two other enzymes (CBH Ia and IIb) also partially lose the CBM during the course of fermentation. The CBM absence in major part of cellobiohydrolases from the NCE-L600 may lead to the lower activity of the crude preparation toward crystalline cellulose.

Table 1

Identification of peptides in the isolated *C. lucknowense* proteins using MALDI-TOF MS/MS

Enzyme	m/z	Peptide ^a	BLAST identification ^b	UniProtKB No.
Protein 60	1133.	HEYGTNIGSR	118 HEYGTNIGSR 127	O94093
kDa	6		(cbh1.2 <i>Humicola grisea</i> – GH7)	
	1829.	MGNQDFYGPGLTVDT	291 LGNTDFYGPGLTVDT 305	Q9UVS8
	9	K	(cbhB <i>Aspergillus niger</i> – GH7)	
Protein 70	1061.	YPANDYYR	127 ANNYYR 132	Q9C1S9
kDa	4		(Avicelase 2 <i>Humicola insolens</i> – GH6)	
	1990.	HYTEAFSPLLNSAGFPAR	367 KYIEAFSPLLNAAGFPA 383	Q872J7
	0		(CBH II <i>Neurospora crassa</i> – GH6)	
	2073.	LWQPTGQQQWGDWCN	381 QPTGQQQWGDWCNV 394	P07987
	5	VK	(CBH II <i>T. reesei</i> – GH6)	

^a Since the MS/MS can not distinguish between Leu and Ile residues (they have the same masses), there may be ambiguity in the appropriate positions of the identified peptides.

^b Residues conserved in the *C. lucknowense* enzymes are shown in **bold**.

Table 2

Specific activities (U mg⁻¹ of protein) of purified cellobiohydrolases from *C. lucknowense* toward different substrates at pH 5.0 and 40°C

Enzyme	Mol. mass (kDa)	Cat. domain designation	CBM presence	Avicel	CMC ^a	Barley β -glucan ^a	<i>p</i> -NP- β -D-cellobioside	<i>p</i> -NP- β -D-lactoside
CBH Ia	65	Cel7A	Yes	0.21	0.1	<0.1	0.021	0.12
CBH Ib	60	Cel7B	No	0.12	0.3	<0.1	0.020	0.09
CBH IIa	43	Cel6A	No	0.08	1.1	2.0	0	0
CBH IIb	70	Cel6B	Yes	0.22	0.2	0.2	0	0

^a Activity was determined at 50°C.

Table 3

Synergism between *C. lucknowense* cellulases in hydrolysis of cotton cellulose (5 mg ml⁻¹) at pH 5.0 and 40°C in the presence of 0.5 U ml⁻¹ of *A. japonicus* BGL. In all cases the CBH concentration was 0.15 mg ml⁻¹, the EG concentration was 0.05 mg ml⁻¹.

Enzyme	Glucose concentration after 140 h, experimental (mg ml ⁻¹)	Glucose concentration after 140 h, theoretical ^a (mg ml ⁻¹)	K_{syn}
CBH Ia	0.81	—	—
CBH IIb	1.18	—	—
EG II	0.64	—	—
EG V	0.70	—	—
EG VI	0.40	—	—
CBH Ia + EG II	4.05	1.45	2.79
CBH Ia + EG V	3.68	1.51	2.44
CBH Ia + EG VI	3.93	1.21	3.25
CBH IIb + EG II	4.72	1.82	2.59
CBH IIb + EG V	3.81	1.88	2.03
CBH IIb + EGVI	4.05	1.58	2.56
CBH Ia + CBH IIb	5.47	1.99	2.75

^a Calculated as a sum of glucose concentrations obtained under the action of individual enzymes.

Table 4

Specific activities (U mg^{-1} of protein) of multienzyme preparations toward different substrates at pH 5.0 and 50°C

Preparation	Protein (mg ml^{-1} or mg g^{-1})	Filter paper	CMC	Xylan	Cellobiose ^a
NCE-L600	45	0.25	12.2	4.8	0.07
<i>C.l.</i> combination #1	1000	1.10	6.6	0	1.05

^a Activity was determined at 40°C.

Table 5

Composition of artificial multienzyme combinations based on purified *C. lucknowense* enzymes and yields of glucose after 72-h hydrolysis of pretreated Douglas fir wood (50 mg ml⁻¹), pH 5.0, 50°C . The total protein concentration in the reaction system was 0.5 mg ml⁻¹, the concentration of each component and glucose yields are given in mg ml⁻¹.

Combination	CBH Ia	CBH Ib	CBH IIb	EG II	EG V	BGL	Xyl II	Glucose yield
#1	0.2	0	0.2	0.08	0	0.02	0	28.8
#2	0.2	0	0.2	0.07	0	0.02	0.01	30.1
#3	0.2	0	0.2	0.04	0.04	0.02	0	37.3
#4	0.1	0.1	0.2	0.03	0.04	0.02	0.01	41.0

CLAIMS

1. An enzyme formulation, said formulation comprising at least one enzyme selected from the group consisting of CBH Ib (SEQ ID NO. 2), CBH IIb (SEQ ID NO. 4), EG II (SEQ ID NO. 10), EG VI (SEQ ID NO. 16), BGL (SEQ ID NO 12), and Xyl II (SEQ ID NO. 18).
2. A method of producing fermentable sugars from lignocellulosic material, said method comprising
 - (a) providing an enzyme formulation, wherein said enzyme formulation comprises at least one enzyme selected from the group consisting of CBH Ib (SEQ ID NO. 2), CBH IIb (SEQ ID NO. 4), EG II (SEQ ID NO. 10), EG VI (SEQ ID NO. 16), BGL (SEQ ID NO 12), and Xyl II (SEQ ID NO. 18);
 - (b) applying said enzyme formulation to lignocellulosic material to produce fermentable sugars.
3. The method according to claim 2, wherein the fermentable sugars comprise at least one sugar from the group consisting of glucose, xylose, arabinose, galactose, mannose, rhamnose, sucrose and fructose.
4. The method according to claim 2, wherein the lignocellulosic material is selected from the group consisting of orchard prunings, chaparral, mill waste, urban wood waste, municipal waste, logging waste, forest thinnings, short-rotation woody crops, industrial waste, wheat straw, oat straw, rice straw, barley straw, rye straw, flax straw, soy hulls, rice hulls, rice straw, corn gluten feed, oat hulls, sugar cane, corn stover, corn stalks, corn cobs, corn husks, prairie grass, gamagrass, foxtail; sugar beet pulp, citrus fruit pulp, seed hulls, cellulosic animal wastes, lawn clippings, cotton, seaweed, trees, shrubs, grasses, wheat, wheat straw, sugar cane bagasse, corn, corn husks, corn kernel, fiber from kernels, products and by-products from wet or dry milling of grains, municipal solid waste, waste paper, yard waste, herbaceous material, agricultural residues, forestry residues, municipal solid wastes, waste paper, pulp, paper mill residues,

branches, bushes, canes, corn, corn husks, energy crops, forests, fruits, flowers, grains, grasses, herbaceous crops, leaves, bark, needles, logs, roots, saplings, shrubs, switch grasses, trees, vegetables, fruit peels, vines, sugar beet pulp, wheat midlings, oat hulls, hard and soft woods, organic waste materials generated from agricultural processes, forestry wood waste, or combinations thereof.

5. A method of producing a fermentation product or a starting material for a fermentation product from a fermentable sugar, wherein said method comprises
 - (a) providing an enzyme formulation, wherein said enzyme formulation comprises at least one enzyme selected from the group consisting of CBH Ib (SEQ ID NO. 2), CBH IIb (SEQ ID NO. 4), EG II (SEQ ID NO. 10), EG VI (SEQ ID NO. 16), BGL (SEQ ID NO 12), and Xyl II (SEQ ID NO. 18);
 - (b) applying said enzyme formulation to lignocellulosic material to produce a fermentable sugar; and
 - (c) fermenting said fermentable sugar to produce a fermentation product.
6. The method according to claim 5, wherein the fermentable sugar is selected from the group consisting of glucose, xylose, arabinose, galactose, mannose, rhamnose, sucrose and fructose.
7. The method according to claim 5, wherein the lignocellulosic material is selected from the group consisting of orchard prunings, chaparral, mill waste, urban wood waste, municipal waste, logging waste, forest thinnings, short-rotation woody crops, industrial waste, wheat straw, oat straw, rice straw, barley straw, rye straw, flax straw, soy hulls, rice hulls, rice straw, corn gluten feed, oat hulls, sugar cane, corn stover, corn stalks, corn cobs, corn husks, prairie grass, gamagrass, foxtail; sugar beet pulp, citrus fruit pulp, seed hulls, cellulosic animal wastes, lawn clippings, cotton, seaweed, trees, shrubs, grasses, wheat, wheat straw, sugar cane bagasse, corn, corn husks, corn kernel, fiber from kernels, products and by-products from wet or dry milling of grains, municipal solid waste, waste paper, yard waste, herbaceous material, agricultural residues,

forestry residues, municipal solid wastes, waste paper, pulp, paper mill residues, branches, bushes, canes, corn, corn husks, energy crops, forests, fruits, flowers, grains, grasses, herbaceous crops, leaves, bark, needles, logs, roots, saplings, shrubs, switch grasses, trees, vegetables, fruit peels, vines, sugar beet pulp, wheat midlings, oat hulls, hard and soft woods, organic waste materials generated from agricultural processes, forestry wood waste, or combinations thereof.

8. The method according to claim 5, wherein said fermentation product is a biofuel.
9. The method according to claim 5, wherein said fermentation product is selected from the group consisting of lactic acid; plastics; specialty chemicals, organic acids, solvents, animal feed supplements, pharmaceuticals, vitamins; amino acids, industrial enzymes, and chemical feedstocks.
10. A method of producing energy from a fermentable sugar, said method comprising
 - (a) providing an enzyme formulation, wherein said enzyme formulation comprises at least one enzyme selected from the group consisting of CBH Ib (SEQ ID NO. 2), CBH IIb (SEQ ID NO. 4), EG II (SEQ ID NO. 10), EG VI (SEQ ID NO. 16), BGL (SEQ ID NO 12), and Xyl II (SEQ ID NO. 18);
 - (b) applying said enzyme formulation to lignocellulosic material to produce a fermentable sugar;
 - (c) fermenting said fermentable sugar to produce a combustible fermentation product;
 - (d) combusting said combustible fermentation product to produce energy.
11. The method according to claim 10, wherein the fermentable sugar is selected from the group consisting of glucose, xylose, arabinose, galactose, mannose, rhamnose, sucrose and fructose.

12. The method according to claim 10, wherein the lignocellulosic material is selected from the group consisting of orchard prunings, chaparral, mill waste, urban wood waste, municipal waste, logging waste, forest thinnings, short-rotation woody crops, industrial waste, wheat straw, oat straw, rice straw, barley straw, rye straw, flax straw, soy hulls, rice hulls, rice straw, corn gluten feed, oat hulls, sugar cane, corn stover, corn stalks, corn cobs, corn husks, prairie grass, gamagrass, foxtail; sugar beet pulp, citrus fruit pulp, seed hulls, cellulosic animal wastes, lawn clippings, cotton, seaweed, trees, shrubs, grasses, wheat, wheat straw, sugar cane bagasse, corn, corn husks, corn kernel, fiber from kernels, products and by-products from wet or dry milling of grains, municipal solid waste, waste paper, yard waste, herbaceous material, agricultural residues, forestry residues, municipal solid wastes, waste paper, pulp, paper mill residues, branches, bushes, canes, corn, corn husks, energy crops, forests, fruits, flowers, grains, grasses, herbaceous crops, leaves, bark, needles, logs, roots, saplings, shrubs, switch grasses, trees, vegetables, fruit peels, vines, sugar beet pulp, wheat midlings, oat hulls, hard and soft woods, organic waste materials generated from agricultural processes, forestry wood waste, or combinations thereof.
13. The method according to claim 10, wherein said combustible fermentation product is an alcohol.
14. The method according to claims 2, 5, or 10, wherein the lignocellulosic material is subjected to a pretreatment prior to being exposed to enzymes.
15. The method according to claim 14, wherein the pretreatment comprises exposing the lignocellulosic biomass to an acid, base, solvent, heat, peroxide, ozone, mechanical shredding, grinding, milling, rapid depressurization, or a combination thereof.

16. The method according to claim 15, wherein said solvent is an acetone/ethanol mixture or organosolv.
17. A mutant *Chrysosporium lucknowense* strain capable of expressing at least one cellobiohydrolase and at least one endo-1,4- β -glucanase at higher levels than the corresponding non-mutant strain under the same conditions;
wherein said at least one cellobiohydrolase is selected from the group consisting of CBH Ia, CBH Ib, CBH IIa and CBH IIb; and
wherein said at least one endo-1,4- β -glucanase is selected from the group consisting of EG II, EG V, and EG VI.
18. The mutant *Chrysosporium lucknowense* strain according to claim 17, wherein said strain is capable of expressing a β -glucosidase and/or a xylanase at higher levels than the corresponding non-mutant strain under the same conditions.
19. The mutant *Chrysosporium lucknowense* strain according to claim 18, wherein said xylanase is Xyl II.
20. The mutant *Chrysosporium lucknowense* strain according to claim 17, wherein said mutant is or is derived from a *Chrysosporium lucknowense* mutant strain selected from the group consisting of *C. lucknowense* strain C1 (VKM F-3500 D), UV13-6 (VKM F-3632 D), NG7C-19 (VKM F-3633 D), and UV18-25 (VKM F-3631 D).
21. A protein exhibiting at least 65% amino acid identity as determined by the BLAST algorithm with the CBH Ib amino acid sequence of SEQ ID NO. 2 or a part thereof having at least 20 contiguous amino acids.
22. The protein according to claim 21, wherein the protein exhibits at least 70% amino acid identity as determined by the BLAST algorithm with the CBH Ib amino acid sequence of SEQ ID NO. 2 or a part thereof having at least 20 contiguous amino acids.

23. The protein according to claim 21, wherein said protein exhibits at least 75% amino acid identity as determined by the BLAST algorithm with the CBH Ib amino acid sequence of SEQ ID NO. 2 or a part thereof having at least 20 contiguous amino acids.
24. A nucleic acid sequence encoding a protein according to claim 21.
25. A nucleic acid sequence encoding a protein according to claim 22.
26. A nucleic acid sequence encoding a protein according to claim 23.
27. A protein exhibiting at least 65% amino acid identity as determined by the BLAST algorithm with the CBH IIb amino acid sequence of SEQ ID NO. 4 or a part thereof having at least 20 contiguous amino acids.
28. The protein according to claim 27, wherein the protein exhibits at least 70% amino acid identity as determined by the BLAST algorithm with the CBH IIb amino acid sequence of SEQ ID NO. 4 or a part thereof having at least 20 contiguous amino acids.
29. The protein according to claim 28, wherein said protein exhibits at least 75% amino acid identity as determined by the BLAST algorithm with the CBH IIb amino acid sequence of SEQ ID NO. 4 or a part thereof having at least 20 contiguous amino acids.
30. A nucleic acid sequence encoding a protein according to claim 27.
31. A nucleic acid sequence encoding a protein according to claim 28.
32. A nucleic acid sequence encoding a protein according to claim 29.

33. A protein exhibiting at least 65% amino acid identity as determined by the BLAST algorithm with the EG VI amino acid sequence of SEQ ID NO. 16 or a part thereof having at least 20 contiguous amino acids.
34. The protein according to claim 33, wherein the protein exhibits at least 70% amino acid identity as determined by the BLAST algorithm with the EG VI amino acid sequence of SEQ ID NO. 16 or a part thereof having at least 20 contiguous amino acids.
35. The protein according to claim 34, wherein said protein exhibits at least 75% amino acid identity as determined by the BLAST algorithm with the EG VI amino acid sequence of SEQ ID NO. 16 or a part thereof having at least 20 contiguous amino acids.
36. A nucleic acid sequence encoding a protein according to claim 33.
37. A nucleic acid sequence encoding a protein according to claim 34.
38. A nucleic acid sequence encoding a protein according to claim 35.
39. A protein exhibiting at least 65% amino acid identity as determined by the BLAST algorithm with the BGL amino acid sequence of SEQ ID NO. 12 or a part thereof having at least 20 contiguous amino acids.
40. The protein according to claim 39, wherein the protein exhibits at least 70% amino acid identity as determined by the BLAST algorithm with the BGL amino acid sequence of SEQ ID NO. 12 or a part thereof having at least 20 contiguous amino acids.
41. The protein according to claim 42, wherein said protein exhibits at least 75% amino acid identity as determined by the BLAST algorithm with the BGL

- amino acid sequence of SEQ ID NO. 12 or a part thereof having at least 20 contiguous amino acids.
42. A nucleic acid sequence encoding a protein according to claim 39.
 43. A nucleic acid sequence encoding a protein according to claim 40.
 44. A nucleic acid sequence encoding a protein according to claim 41.
 45. A protein exhibiting at least 65% amino acid identity as determined by the BLAST algorithm with the Xyl II amino acid sequence of SEQ ID NO. 18 or a part thereof having at least 20 contiguous amino acids.
 46. The protein according to claim 45, wherein the protein exhibits at least 70% amino acid identity as determined by the BLAST algorithm with the Xyl II amino acid sequence of SEQ ID NO. 18 or a part thereof having at least 20 contiguous amino acids.
 47. The protein according to claim 46, wherein said protein exhibits at least 75% amino acid identity as determined by the BLAST algorithm with the Xyl II amino acid sequence of SEQ ID NO. 18 or a part thereof having at least 20 contiguous amino acids.
 48. A nucleic acid sequence encoding a protein according to claim 45.
 49. A nucleic acid sequence encoding a protein according to claim 46.
 50. A nucleic acid sequence encoding a protein according to claim 47.
 51. A nucleic acid sequence having at least 80% homology with the nucleic acid sequence encoding CBH Ib as shown in SEQ ID NO. 1.

52. A nucleic acid sequence having at least 80% homology with the nucleic acid sequence encoding CBH IIb as shown in SEQ ID NO. 3.
53. A nucleic acid sequence having at least 80% homology with the nucleic acid sequence encoding EG VI as shown in SEQ ID NO. 15.
54. A nucleic acid sequence having at least 80% homology with the nucleic acid sequence encoding BGL as shown in SEQ ID NO. 11.
55. A nucleic acid sequence having at least 80% homology with the nucleic acid encoding Xyl II as shown in SEQ ID NO. 17.
56. A method for degrading a lignocellulosic material to fermentable sugars, said method comprising
- contacting the lignocellulosic material with an effective amount of a multi-enzyme product derived from a microorganism, to produce at least one fermentable sugar
- wherein at least one of enzyme in the multi-enzyme product is selected from the group consisting of CBH Ia (SEQ ID NO. 6), CBH Ib (SEQ ID NO. 2), CBH IIa (SEQ ID NO. 8), CBH IIb (SEQ ID NO. 4), EG II (SEQ ID NO. 10), EG V (SEQ ID NO. 14), EG VI (SEQ ID NO. 16), BGL (SEQ ID NO. 12), and Xyl II (SEQ ID NO. 18).
57. A microorganism or plant capable of expressing one or more of an enzyme selected from the group consisting of CBH Ia (SEQ ID NO. 6), CBH Ib (SEQ ID NO. 2), CBH IIa (SEQ ID NO. 8), CBH IIb (SEQ ID NO. 4), EG II (SEQ ID NO. 10), EG V (SEQ ID NO. 14), EG VI (SEQ ID NO. 16), BGL (SEQ ID NO. 12), and Xyl II (SEQ ID NO. 18).
58. The microorganism according to claim 57, wherein said microorganism is a fungus.

59. The microorganism according to claim 58, wherein said fungus is selected from the group consisting of *Talaromyces*, *Aspergillus*, *Trichoderma*, *Neurospora*, *Penicillium*, *Fusarium*, *Humicola*, *Myceliophthora*, *Corynascus*, *Chaetomium*, *Tolypocladium*, *Thielavia*, *Acremonium*, *Sporotrichum*, *Thermoascus*, and *Chrysosporium*.
60. The microorganism according to claim 59, wherein said fungus is *Chrysosporium lucknowense*.
61. The microorganism according to claim 60, wherein said *Chrysosporium* is a strain selected from the group consisting of the wild-type strain, Acc. No. VKM F-3500 D; C1 UV13-6, Acc. No. VKM F-3632 D; strain C1 NG7C-19, Acc. No. VKM F-3633 D; and C1 UV18-25, Acc. No. VKM F-3631 D.
62. A nucleic acid sequence having at least 80% homology with the nucleic acid sequence encoding EG II as shown in SEQ ID NO. 9.

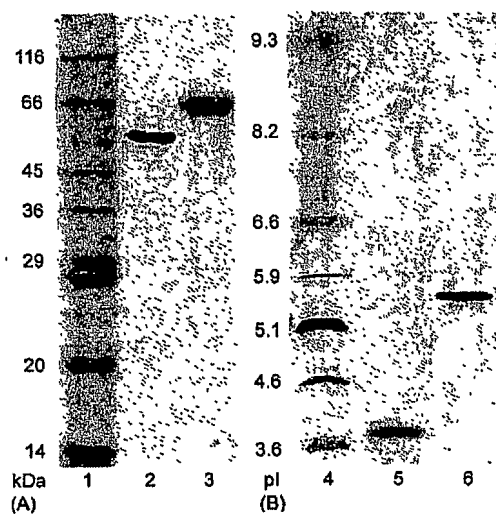


FIGURE 1

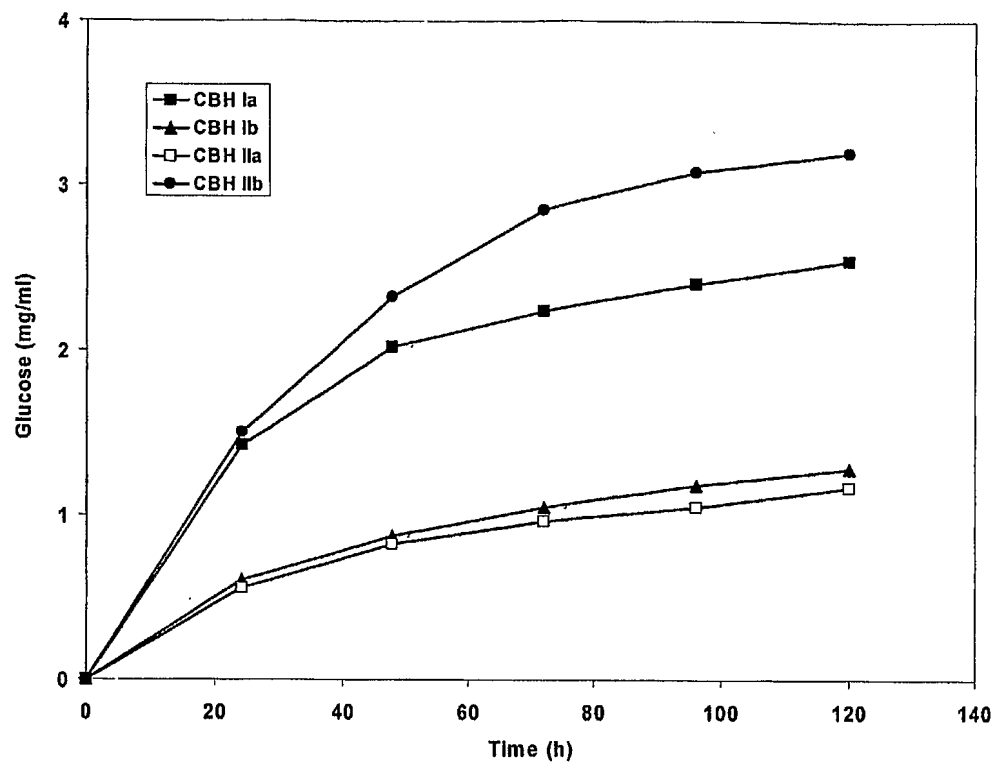


FIGURE 2

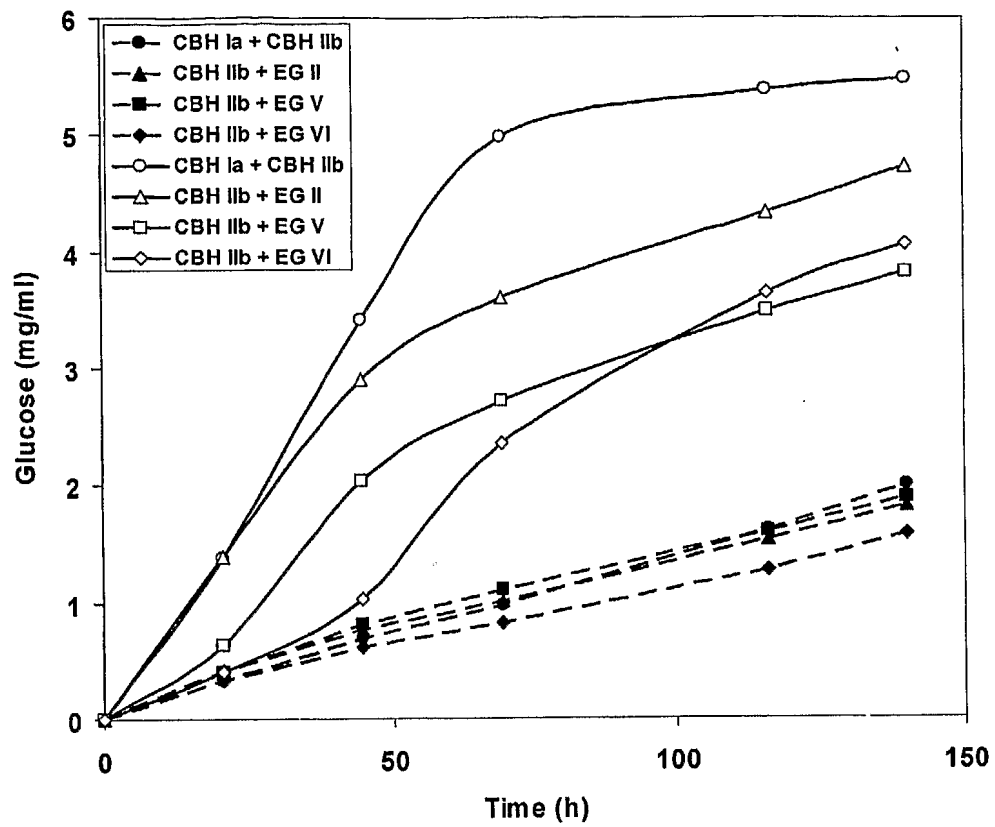


FIGURE 3

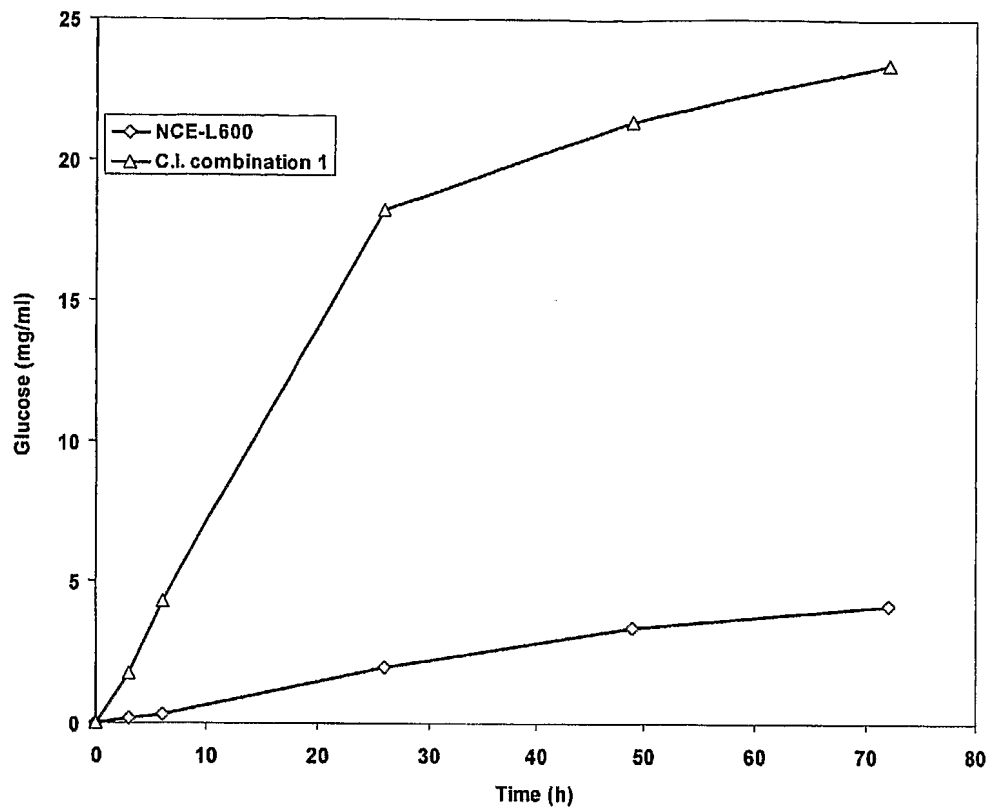


FIGURE 4

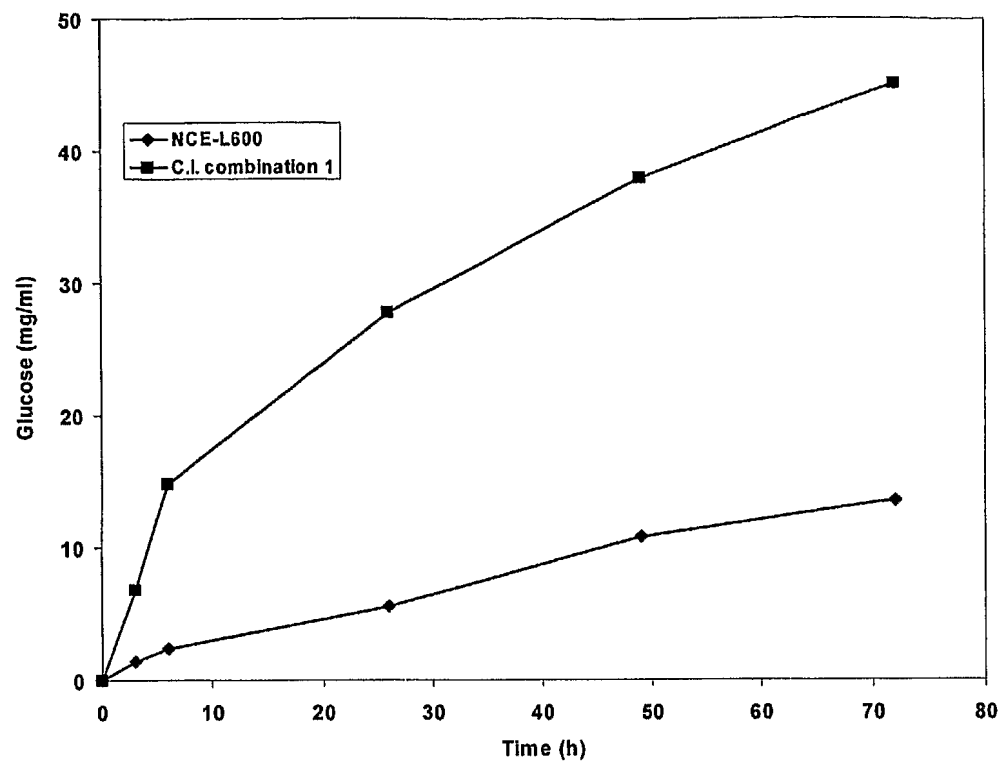


FIGURE 5

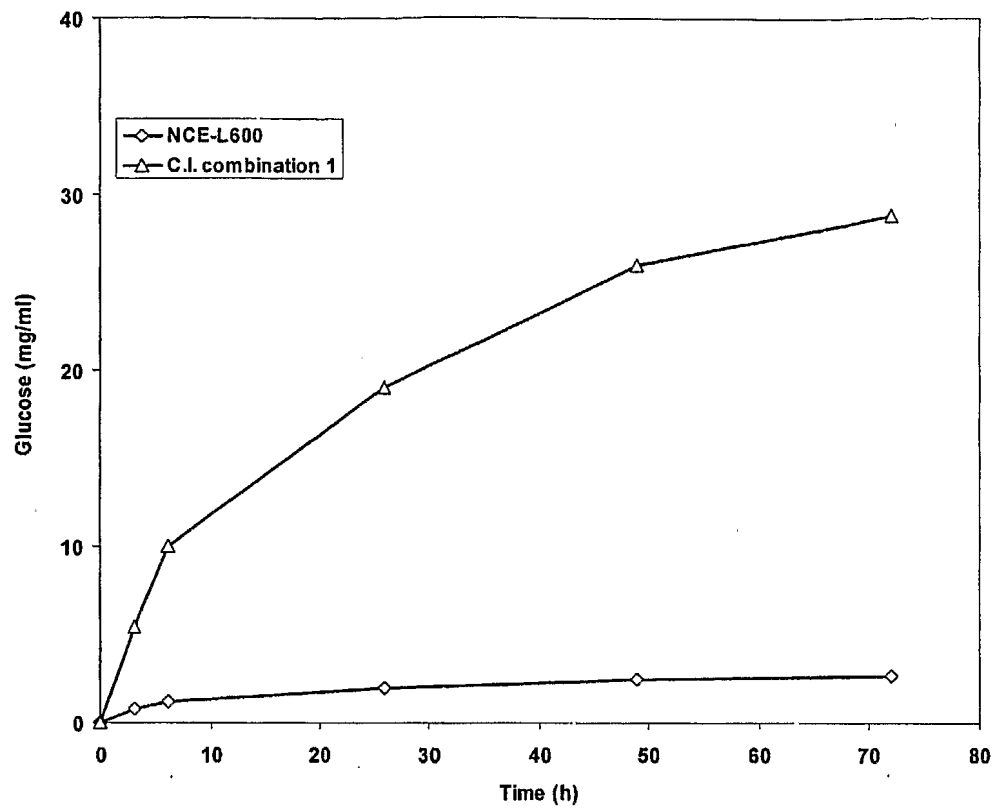


FIGURE 6

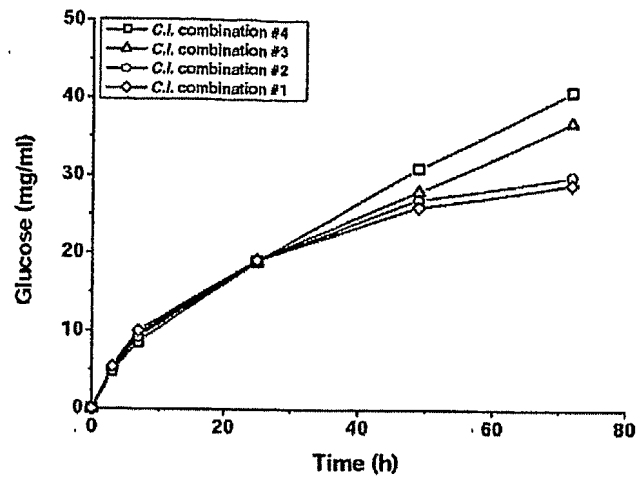


FIGURE 7

Serial No. PCT/US2006/027347Docket No. 3123-4008PC

FIGURE 8

Translation of Contig 2370 14521-20840 cbh2(1-6360)

```

1      ctcagattctagggtagggcgggagcagaggcgaattgggtgtagaatatgaggag
61     ctagggtgtgttaaaactcaaagaactcttgcctcttgccttagtctctctcctgggaaa
121    aggggggtttttccgaagcggcgctatacgaagccagaggtactttccttgccttggat
181    ggcccttgcaccgcttcttgccttcccggttgcctcaattgcgacgttgcgggcaacctagg
241    tcctaataattaggttagatatttcggtagaggtagtttaattatgcttcagtagagaaat
301    cgttgcctccacgtctcgcacacctgcgaaacttcgccacattgaagatagcattgtctg
361    agttgattttaacctttccagagacgatataatagtgaagtcttcttgcctcggaatca
421    tcgacattcggttttcccttaattatagaagtattcgcccaacggaacggggcccgga
481    gcaggttgaaccgcgcaaaacctcaaccgagtcacctcgctccatggttgcctatggaat
541    cagggtccgaatcccgtagatcagtcagtcctggtggctatggacggcggttgcctacggc
601    cagtcgctcccggttgcctggtgggttgatcaacaggaggaagagatctgagatcgaaacta
661    caccattgattttatcgacgcataatcaagtttaataaaaaacaaacagcgttgcctggtg
721    ctaccaccgaatgcgagatccgggctagcccgcggaagatgatggccacagatctagcg
781    tcatgtatgatttatcctatgcctctatctcgtatctgcctcggttgcgaacacctg
841    accgagagacgactcgacaacctgacacttggcaaaagacatttcgggtgacagcgggag
901    aactccagcgaggaagtgcggcagagatgcggatgagaagacaaacggcgagacgtgcgg
961    cgttggctctcccaaatcgagcgcgactcttcggttgcgaatctccgggataaaatcc
1021   cagcggggggtcagctcagcttgcctggtggagggcgaggacagccatccagcgaggccat
1081   ggaagagaaacaaatcttgggggttagcgaccgagcgaagggggggggggaagcgggag
1141   ggggaagaagtgggtatagagcacgcaccggaaacgcatttgggccccttgccaaacaaaca
1201   ccacaccccgctcctgggagcaagacatccaggatgcaaccagtaggggatgccaaga
1261   agcatctacggcaccatctgcggcgctcgcctgttagagtcccggaaccccgcaatgg
1321   ggccgtgctgggcccctgcccgcgaatgctggcgagcgccatcaacaacattgctcgggg
1381   agggggccgattttatgattagcaaaaaaacaattaaattacccttccattccagcaga
1441   gcttctcctccacggcgggcgggacgcgttgcggacggcggtacactacaacggcgggg
1501   ctccagctcctcgctgctgggctgcagatcacgaccgggaagagaaatgatcgcggtctga
1561   cgccgggtacggagtagcagccgcaaccacagccgagtgaccgtgatctcaatgcg
1621   ttcaagcaacacagcaacacacctggacgagtcctcctccctaccacccctccccccc
1681   tgccctggcgcgaaacggggcgcgtaaccacagatttctactccgtagacaccccaatc
1741   tattccgctggcgctgcggcagtcggtggcggtccggcgaagactctcggtgcagatac
1801   cgcgacgaattcggttaaccgttggctgatcaattccaagtcaaggagagagtggtatg
1861   gaaagtgcggctcagtttccactgccccgcagcgaggttccggatctggacagcagtc
1921   ttccgaatcttggcagagactcatgataataaaaaaggcaaatgaggcgggcgcttgg
1981   acaggtccattctccaccgctcaaccagcctccaattcctcagaagtcgtgtgctctct
2041   cgcagtcgcagtcgaagATGAAGCAGTACCTCCAGTACCTCGCGGCGACCCCTGCCCTGGT
      M K Q Y L Q Y L A A T L P L V
2101   GGGCCTGGCCACGGCCACGAGCGGGTAACCTGCAGACCGAGACTCACCCCAAGCTCAC
      G L A T A Q Q A G N L Q T E T H P K L T
2161   TTGGTTCGAAGTGCACGGCCCCGGGATCCTGCCAACAGGTCAACGGCGAGGTGCTCATCGA
      W S K C T A P G S C Q Q V N G E V V I D
2221   CTCCAATGGCGCTGGGTGCACGACGAGAACGCGCAGAAGTGTACGACGGCAACCAAGT
      S N W R W V H D E N A Q N C Y D G N Q W
2281   GACCAACGCTTCAGCTCTGCCACCGACTGCGCCGAGAATTGCGCGCTCGAGGGTGCCGA
      T N A C S S A T D C A E N C A L E G A D
2341   CTACCAGGGACCTATGGCGCCTCGACCGCGCAATGCCCTGACGCTCACCTTCGTCAC
      Y Q G G T Y G A S T S G N A L T L T F V T
2401   TAAGCAGGATACGGCACCAACATTGGTTCGCGCCTCTACCTCATGAACGGCGCAACAA
      K H E Y G T N I G S R L Y L M N G A N K
2461   GTACCAGATGTTACCCCTCAAGGGCAACGAGCTGGCCTTCGACGTCGACCTCTCGGCCGT
      Y Q M F T L K G N E L A F D V D L S A V
2521   CGAGTGGCGCCTCAACAGCGCCCTCTACTTCGTGGCCATGGAGGAGGATGGCGGTGTGTC
      E C G L N S A L Y F V A M E E D G G V S
2581   GAGCTACCCGACCAACACGGCCGCTGCTAAGTTCGGCACTGGGgtaagttcaacgacccg
      S Y P T N T A G A K F G T G
2641   agacgggtgcccattattatctgctgcgaaaacggacgggtccccttttgtaactaccctc
2701   ctccaaacagTACTGCGACGCCAATGCGCACGCGACCTCAAGTTCGTCGGCGGCAAGGG
      Y C D A Q C A R D L K F V G G K G
2761   CAACATCGAGGGCTGGAAGCCGTCCACCAACGATGCCAATGCCGGTGTGCGTCTTATGG
      N I E G W K P S T N D A N A G V G P Y G
2821   CGGGTGTGCGCTGAGATCGACGCTGTGtaagtttgcctggcgagcaatggtatat
      G C C A E I D V W
2881   tagctcgaattgggttcgctgacccctctcttaccagGGAGTCGAACAAGTATGC
      E S N K Y A
2941   TTTCGCTTTCACCCCGCACGGTTGCGAGAACCCTAAATACCAGCTGTGCGAGACCACCAA
      F A F T P H G C E N P K Y H V C E T T N

```

Serial No. PCT/US2006/027347Docket No. 3123-4008PC

(FIGURE 8, continued)

```
3001 CTGCGGTGGCACCTACTCCGAGGACCGCTTCGCTGGTGACTGCGATGCCAACGGCTGCGA
    C G G T Y S E D R F A G D C D A N G C D
3061 CTACAACCCCTACCGCATGGGCAACCAGGACTTCTACGGTCCCGGCTTGACGGTCGATAC
    Y N P Y R M G N Q D F Y G P G L T V D T
3121 CAGCAAGAAGTTCACgtgagtacacggtgcttgaagcccccctcccccccccccccaaaa
    S K K F T
3181 aaaaaaagaaaaaagaagtcaaatgattgatgctaaccaaatcaaaataacagCGTCGTCA
    V V
3241 GCCAGTTCGAGGAGAACAAAGCTCACCCAGTTCTTCGTCAGGACGGCAAGAAGATTGAGA
    S Q F E E N K L T Q F F V Q D G K K I E
3301 TCCCGGGCCCCAAGGTGAGGGCATCGATGCGGACAGCGCCGCTATCACCCCTGAGCTGT
    I P G P K V E G I D A D S A A I T P E L
3361 GCAGTGCCTGTTCAAGGCCTTCGATGACCGTGACCGCTTCTCGGAGGTGGCGGCTTCG
    C S A L F K A F D D R D R F S E V G G F
3421 ATGCGCATCAACCGGCCCTCAGCACTCCCATGGTCTCGTCATGTCCATCTGGGATGATg
    D A I N T A L S T P M V L V M S I W D D

3481 tacgttacctaaccccccccccctttttttttcccgcttctctccccgaaactgccacta
3541 cttatatacgtcccgcgctccatgatgcttaccttttctccttccagCACTACGCCAATAT
    H Y A N M
3601 GCTCTGGCTCGACTCGAGCTACCCCCCTGAGAAGGCTGGCCAGCCTGGCGGTGACCGTGG
    L W L D S S Y P P E K A G Q P G G D R G
3661 CCCGTGTCTCAGGACTCTGGCGTCCCGGCCGACGTTGAGGCTCAGTACCCTAATGCGTG
    P C P Q D S G V P A D V E A Q Y P N A *
3721 agtcgaaaccgttaaatgtcgggcaaaaaaagatcgctcaagctaacgaaataaatga
3781 ttagCAAGGTCACTCTGGTCCAACATCCGCTTCGGCCCCATCGGCTCGACTGTCAACGTCT
    K V I W S N I R F G P I G S T V N V
3841 AAactgcaacctgaccgggcccctttctctccacccccacccctctcaagttctctctggt
3901 ggagccctcgtgtccttcttttcttaggttcgcaaccttttagcttgtgtatcgtaggg
3961 tcatgtgtacatacacaaaaacttaacatctgctaccaagatcttggcgcttggccagg
4021 tcttctcaaacctcgaagcactgagccttctgctccgagtgaaagtaggatgactattta
4081 cgttgcgaagactacgcggtaaaggggacggagcagacctgccacagatattcgtttggtt
4141 gcttgatttatagcagagtcgcaacgttagacatggccctgaaggtgccaacctagata
4201 ttgagccagcaacacatggtagaatagccagcaaaaggacgcttgttcgcttgatctcgaa
4261 taccgcgacgcaatgtatcgtaagaatgtgaactaaagggaaacgagagcatagggaaa
4321 cgtcaatgtggctgaataacagagtttaataacttaataagaagaattagcatgccaaga
4381 ttgagccagcaacacatggtagaatagccagcaaaaggacgcttgttcgcttgatctcgaa
4441 ccgtccaacctgattcgaaggaggagggaaggttgaagaataccggcaataattactcg
4501 aggttccatagccctgcagagttcaatataataataaaggcaccacccgatgattccgca
4561 attataagcataaataagctcgggccccacacgctgaccttcaacctccatgtgtatata
4621 atctgtacctcgttattgtcgaatcgctattccgatagcgaaggtctggcactcatcaga
4801 taccgtgacatcgattgagatttggcggggccacgggtagtaagcgatgagttggtcatc
4681 aattatcaacaatgctcgaatcgaatcagcgataatcagcctatcaaccgcgaatacatacgcg
4741 catcaacgaattgtccatcatgcacgtagcttgtcggcagtgccgcataccctccagagc
4861 atcatagccgggtagaagaagctcgctttcagccgtcccagagtcggagatgaggtagca
4921 agccttcaagaccagttatatgtgacccgggtaaaataacttggtgagatgcaatggcggt
4981 agcttccgggcacttataagctttactagatattatctcaaggtttcttttgaactcctc
5041 ctagacatttactataaactaccgagcttcaatgctagacgacctccttctgttaaatag
5101 tcttttcttctaagagcatctgccttttttcccttaggcttagaggatagggccccctcc
5161 atcttgcgtcgacggccttagccttggggagtaattattgggtatccgcgtacctgtttcc
5221 cagacagccgaagtttcgacgacaaagtaattattgacgaataaccaccgccatagcta
5281 ttccgagtggttgagccccgaaaaacatcgcttaccgcatcgccatcccagacgacagagg
5341 gcgactttgatgtcttgctccagatcgccgcacctaacgggtggggtgggctgggtatcg
5401 tatgggacggcatcatggtcaacaacccccctgacgggggtgttgggccaatggaaacacca
5461 ccggtgtctcgagccggatcgcaaggtgaagccgaaggacaaatgacgatgagactttc
5521 tttcttttttttttttttttttttttaatttcttttttaagcgtaatagaagagctaca
5581 tatctgtggttcggttcctcaatttcagcgacctctccaccgaagcatcgtaaatgaaga
5641 gttgtcggaacaaagggtgtcagaagctatagagcttctaaggatattagccacataca
5701 tgccatagctgtataaggctatttaacgctttggccagctcctttgtctataaatattag
5761 tcgttttgtctcctttgtagataattttaacaaggcactcttttctttatagccacc
5821 tactatagactgcttcaacgctcccggaagcttat tactacgttcggcagttataagcc
5881 tggcgcttgactactcctctgcgcagctatctttaatatagtagtagcttcttctatt
5941 acgaactctcttaccctgctttaatacgttttcgacgacgtgtctattatatactaaatc
6001 ctagtgcagacttctatatgccttactaggcctagttcttagaacttgtagtatattaaa
6061 ctatagttataggctaaatttgctagtatatagagatttggtaaccttaatagtaattat
6121 aaactagatctagaagttttatagtcctaacctataaataagctagagataaccttatt
6181 ttagcttcttaggagtaattcctagaaggagattacctttaatatctatagatttgata
6241 ccttctaataatagctatcatagctaaatttatataattataagattccttttataaaaat
```

Serial No. PCT/US2006/027347

Docket No. 3123-4008PC

6301	attatataactatagatattagtaagtagataggatagctataataactagctagtatat
------	--

Serial No. PCT/US2006/027347Docket No. 3123-4008PC

FIGURE 9

SEQ ID NOs: 3 and 4 cbh4 gene encoding CBH IIb

```

1      cgcgaagtgaatatgtaattactcaatggaagttctcgaaacggagtcagaaaatgatgt
61     gggtctgtgggaatgcggcaagagggcgacgttgccgtgaatgctgaacattcccgcctc
121    ttcttcttctcgtcttcttcttcttcttcttcttctcggtcgcggtatggttgacggccagcg
181    tgcgcacggctgctgtttatcgagcgtcggtacgtctagccaacatcccgtagacacgac
241    gaccaagcgtcttgagaatgcaacaacgtctcggaacctggcacgcatcttccgccgcag
301    gtcggcagacgcgcctgggcaataccacccctgtccaggccctttccccgcaggcagag
361    ccgcgctcttcttcttcatgggtatttcaggaacgtggcttccgagattctcgctgttctc
421    cccagtcacactgcccagaccgtaacccgggtccaccaccgcggactgtccgaaaaacctg
481    gttcgcccgagattaatatgctatttccggactaagtgcacaacacacaagcacccttc
541    cgcctcgcgctctagaatctgctttctaaacccgggttctcgggcccttccctttcgcgacg
601    cctccgctctccttaccaggcaccatccgcaataggttaaggtagccaaccgttttggagc
661    gtgattctgccaaggaccgcatccttgcaattcgccatctggtcaaggaccctcttccc
721    gctccattctgggtggctctatcgggacggcggttccccatggctctccaggagagtgtgt
781    gcgagtctggagagccggggttggegtcacgatgctgcccacctagggccggccagcccg
841    gcactgcgctcccgttgatccgtctatccccgtcaagagcaccagccccggcgctcgtga
901    attttcgacttggtcgacttgctacaggtgataaagaggatgcacgccgccctcgatcgg
961    cctgtgtgggttctctccctcgtgccaaaccactcccacctcccgcgccgagatagttgc
1021   ttgtttcgctccgtgagagggacacacaccaATGGCCAAGAAGCTTTTCATCACCGCCGC
                                     M A K K L F I T A A

1081   GCTTGCGGCTGCCGTGTTGGCGGCCCCCGTCATTGAGGAGCGCCAGAACTGCGGCGCTGT
361     L A A A V L A A P V I E E R Q N C G A V

1141   gtggttaagaaagcccgggtccgagtctcccatgattttctcgctcgagtaatggcataaggg

1201   ccaccccttcgactgaccgtgagaatcgatcaaatccagGACTCAATGCGGCGGTAACGG
                                     T Q C G G N G

1261   GTGGCAAGGTCCCACATGCTGCGCCTCGGGCTCGACCTGCGTTGCGCAGAACGAGTGGTA
       W Q G P T C C A S G S T C V A Q N E W Y

```

Serial No. PCT/US2006/027347Docket No. 3123-4008PC

FIGURE 9 (CONT'D)

SEQ ID NOs: 3 and 4 cbh4 gene encoding CBH IIb

1321 CTCTCAGTGCCTGCCCAACAGCCAGGTGACGAGTTCACCACTCCGTCGTCGACTTCCAC
S Q C L P N S Q V T S S T T P S S T S T

1381 CTCGCAGCGCAGCACCAGCACCTCCAGCAGCACCACCAGGAGCGGCAGCTCCTCCTCCTC
S Q R S T S T S S S T T R S G S S S S S

1441 CTCCACCACGCCCCGCGTCTCCAGCCCCGTGACCAGCATTCGCGGCGGTGCGACCTC
S T T P P P V S S P V T S I P G G A T S

1501 CACGGCGAGCTACTCTGGCAACCCCTTCTCGGGCGTCCGGCTCTTCGCCAACGACTACTA
T A S Y S G N P F S G V R L F ██████████

1561 CAGGTCCGAGGTCCACAATCTCGCCATTCTAGCATGACTGGTACTCTGGCGGCCAAGGC
██████ S E V H N L A I P S M T G T L A A K A

1621 TTCCGCGCTCGCCGAAGTCCCTAGCTTCCAGTGGCTCGACCGGAACGTCACCATCGACAC
S A V A E V P S F Q W L D R N V T I D T

1681 CCTGATGGTCCAGACTCTGTCCCAGGTCCGGGCTCTCAATAAGGCCGGTGCCAACTCTCC
L M V Q T L S Q V R A L N K A G A N P P

1741 CTATGCTGgtgagttacatggcgacttgcttctcgtccccctacctttcttgacgggatc
Y A

1801 ggttacctgacctggaggcaaaacaacaacagCCCAACTCGTCGTCTACGACCTCCCCGA
A Q L V V Y D L P D

1861 CCGTGACTGTGCCCGCGCTGCGTCCAACGGCGAGTTTTTCGATTGCAAACGGCGGCGCCGC
R D C A A A A S N G E F S I A N G G A A

1921 CAACTACAGGAGCTACATCGACGCTATCCGCAAGCACATCATTGAGTACTCGGACATCCG

Serial No. PCT/US2006/027347Docket No. 3123-4008PC

FIGURE 9 (CONT'D)

SEQ ID NOs: 3 and 4 cbh4 gene encoding CBH IIb

 N Y R S Y I D A I R K H I I E Y S D I R

1981 GATCATCCTGGTTATCGAGCCCGACTCGATGGCCAACATGGTGACCAACATGAACGTGGC
 I I L V I E P D S M A N M V T N M N V A

2041 CAAGTGCAGCAACGCCGCGTCGACGTACCACGAGTTGACCGTGTACGCGCTCAAGCAGCT
 K C S N A A S T Y H E L T V Y A L K Q L

2101 GAACCTGCCCCAACGTCGCCATGTATCTCGACGCCGGCCACGCCGGCTGGCTCGGCTGGCC
 N L P N V A M Y L D A G H A G W L G W P

2161 CGCCAACATCCAGCCCGCCCGGAGCTGTTTGCCGGCATCTACAATGATGCCGGCAAGCC
 A N I Q P A A E L F A G I Y N D A G K P

2221 GGCTGCCGTCCGCGGCCTGGCCACTAACGTCGCCAACTACAACGCCTGGAGCATCGCTTC
 A A V R G L A T N V A N Y N A W S I A S

2281 GGCCCCGTCGTACACGTCGCCTAACCTAACTACGACGAGAAGCAC'TACATCGAGGCCTT
 A P S Y T S P N P N Y D E K [REDACTED]

2341 CAGCCCGCTCTTGAACCTCGGCCGGCTTCCCCGCACGCTTCATTGTCGACACTGGCCGCAA
 [REDACTED] F I V D T G R N

2401 CGGCAAACAACCTACCGgtatgttttttttttcttttgtctctgtcccccccttttctccc
 G K [REDACTED]

2461 ccttcagttggcggtccacaaggtctcttagtctgtgcttcatctgtgaccaacctcccccc

2521 ccccggcaccgcccacaaccgtttgactctatactcttgggaatgggcgccgaaactgac

2581 cgttccacagGCCAACAACAGTGGGGTGACTGGTGCAATGTCAAGGGCACCGGCTTTGGC

Serial No. PCT/US2006/027347Docket No. 3123-4008PC

FIGURE 9 (CONT'D)

SEQ ID NOs: 3 and 4 cbh4 gene encoding CBH IIb

G T G F G

2641 GTGCGCCCGACGGCCAACACGGGCCACGAGCTGGTCGATGCCTTTGTCTGGGTCAAGCCC
V R P T A N T G H E L V D A F V W V K P

2701 GGCGGCGAGTCCGACGGCACAAGCGACACCAGCGCCCGCCGCTACGACTACCACTGCGGC
G G E S D G T S D T S A A R Y D Y H C G

2761 CTGTCCGATGCCCTGCAGCCTGCCCCCGAGGCTGGACAGTGGTTCCAGGCCTACTTCGAG
L S D A L Q P A P E A G Q W F Q A Y F E

2821 CAGCTGCTCACCAACGCCAACCCGCCCTTCTAAacctcgtcataaagagagagagatggc
Q L L T N A N P P F *

2881 gggcatgggcctgattgggttcattgaccatgcggtcttctctgggggtacatattttacc
2941 tacctacctataaataaggcggcctatcggtctctcgcttcggtttattaggtacttggtc
3001 ttgtacatactttgtttatacatagcaggttagcatccactattcggttcgacaaagcg
3061 gaactttccagaaaaaaaaaaggttgatcataattagtccttaggcttcgattccttggtgc
3121 ctttcttttttggtataaaaaaaatttttttttgaggcatgattaccttaggtacgttcgtc
3181 gttgtattgggtccccctgcattttggcgcgagagcagctcagccccctgcaaaccctca
3241 acgggcggttcaattccctccactcgggtcttcagcgagaccagccgtccagagtatccca
3301 gcgtgtagttgccccacgaaccagtcgtcctcgtaagcctcgtaaaagtgtccaagagca
3361 gtatagaagcaaacgacctccgtcaaaagtctggcaccatgcgatcggttggtcctccccg
3421 tgcgccccgcctcgttaggacttctcatccacgccaaggagcagtgagggcggtcggtc
3481 gtcgccccgcgggtgcgccttgaagttgtaccattcgtccttcagacgcgctccagctgc
3541 gcctgcttgggttcctgcggttcctgcggttcctgcgctggccggtcggcgcgcggtct
3601 tggtcacacgccccgcagcgacatgactgggtgttccgggtcgagcagcttgacgagcccg
3661 acctgggggttcgggtgggtgtcgaacacggcgccaatgaggtggccgtaccattcggtat
3721 gactgcatggcgaagctggcgagtgtagccgacgatcccgccgcccgcctggacgaaa
3781 ccccgagggcgcccagctgcgcgcggtccaggaactcgcccgagcactgcaggaggacg
3841 atgacgcgatacgccgagagggagccggggctgaacacggcgggatcctcgctgtcgtcc

Serial No. PCT/US2006/027347Docket No. 3123-4008PC

FIGURE 10

SEQ ID NOs: 5 and 6 cbh1 gene encoding CBH Ia

```

                                     ATGTACGCCAAGTTCGCGACC 1800
                                     M Y A K F A T
CTCGCCGCCCTTGTGGCTGGCGCCGCTGCTCAGAACGCCTGCACTCTGACCGCTGAGAAC 1860
L A A L V A G A A A Q N A C T L T A E N
CACCCTCGCTGACGTGGTCCAAGTGCACGTCTGGCGGCAGCTGCACCAGCGTCCAGGGT 1920
H P S L T W S K C T S G G S C T S V Q G
TCCATCACCATCGACGCCAACTGGCGGTGGACTCACCGGACCGATAGCGCCACCAACTGC 1980
S I T I D A N W R W T H R T D S A T N C
TACGAGGGCAACAAGTGGGATACTTCGTACTGCAGCGATGGTCCTTCTTGCGCCCTCCAAG 2040
Y E G N K W D T S Y C S D G P S C A S K
TGTCGCATCGACGGCGCTGACTACTCGAGCACCTATGGCATCACCAGAGCGGTAAGTCC 2100
C C I D G A D Y S S T Y G I T T S G N S
CTGAACCTCAAGTTCGTACCAAGGGCCAGTACTCGACCAACATCGGCTCGCGTACCTAC 2160
L N L K F V T K G Q Y S T N I G S R T Y
CTGATGGAGAGCGACACCAAGTACCAGAgtaagttcctctcgaccccgccgcccgggaga 2220
L M E S D T K Y Q M
tgatggcgcccagcccgtgaocggaatgacacaGTGTTCCAGCTCCTCGGCAACGAGTT 2280
                                     F Q L L G N E F
CACCTTCGATGTCGACGTCTCCAACCTCGGCTGCGGCCCTCAATGGCGCCCTCTACTTCGT 2340
T F D V D V S N L G C G L N G A L Y F V
GTCCATGGATGCCGATGGTGGCATGTCCAAGTACTCGGGCAACAAGGCAGGTGCCAAGTA 2400
S M D A D G G M S K Y S G N K A G A K Y
CGGTACCGGCTACTGTGATTCTCAGTGCCCCCGCGACCTCAAGTTCATCAACGGCGAGGC 2460
G T G Y C D S Q C P R D L K F I N G E A
CAACGTAGAGAACTGGCAGAGCTCGACCAACGATGCCAACGCCGCGACGGGCAAGTACGG 2520
N V E N W Q S S T N D A N A G T G K Y G
CAGCTGCTGCTCCGAGATGGACGTCTGGGAGGCCAACACATGGCCGCCGCCCTTCACTCC 2580
S C C S E M D V W E A N N M A A A F T P
CCACCCTTGCNCCGTGATCGGCCAGTCGCGCTGCGAGGGCGACTCGTGCGGCGGTACCTA 2640
H P C ? V I G Q S R C E G D S C G G T Y
CAGCACCGACCGCTATGCCGGCATCTGCGACCCCGACGGATGCGACTTCAACTCGTACCG 2700
S T D R Y A G I C D P D G C D F N S Y R
CCAGGGCAACAAGACCTTCTACGGCAAGGGCATGACGGTCGACACGACCAAGAAGATCAC 2760
Q G N K T F Y G K G M T V D T T K K I T
```

Serial No. PCT/US2006/027347Docket No. 3123-4008PC

FIGURE 10 (CONT'D)

SEQ ID NOs: 5 and 6 cbh1 gene encoding CBH Ia

```
GGTCGTCACCCAGTTCCTCAAGAACTCGGCCGGCGAGCTCTCCGAGATCAAGCGGTTCTA 2820
V V T Q F L K N S A G E L S E I K R F Y

CGTCCAGAACGGCAAGGTCATCCCCAACTCCGAGTCCACCATCCCGGGCGTCGAGGGCAA 2880
V Q N G K V I P N S E S T I P G V E G N

CTCCATCACCCAGGACTGGTGCGACCGCCAGAAGGCCGCCTTCGGCGACGTGACCGACTT 2940
S I T Q D W C D R Q K A A F G D V T D ?

NCAGGACAAGGGCGGCATGGTCCAGATGGGCAAGGCCCTCGCGGGGCCCATGGTCCTCGT 3000
Q D K G G M V Q M G K A L A G P M V L V

CATGTCCATCTGGGACGACCACGCCGTCAACATGCTCTGGCTCGACTCCACCTGGCCCAT 3060
M S I W D D H A V N M L W L D S T W P I

CGACGGCGCGCGCAAGCCGGGCGCCGAGCGCGGTGCCTGCCCCACCACCTCGGGCGTCCC 3120
D G A G K P G A E R G A C P T T S G V P

CGCTGAGGTCGAGGCCGAGGCCCACTCCAACGTCATCTTCTCCAACATCCGCTTCGG 3180
A E V E A E A P N S N V I F S N I R F G

CCCCATCGGCTCCACCGTCTCCGGCCTGCCGACGGCGGCAGCGGCAACCCCAACCCGCC 3240
P I G S T V S G L P D G G S G N P N P P

CGTCAGCTCGTCCACCCCGGTCCCCTCCTCGTCCACCACATCCTCCGGTTCCTCCGGCCC 3300
V S S S T P V P S S S T T S S G S S G P

GACTGGCGGCACGGGTGTCGCTAAGCACTATGAGCAATGCGGAGGAATCGGGTTCACTGG 3360
T G G T G V A K H Y E Q C G G I G F T G

CCCTACCCAGTGCGAGAGCCCCTACACTTGCACCAAGCTGAATGACTGGTACTCGCAGTG 3420
P T Q C E S P Y T C T K L N D W Y S Q C

CCTGTAA
L *
```

Serial No. PCT/US2006/027347Docket No. 3123-4008PC

FIGURE 11

SEQ ID NOs: 7 and 8 eg6 gene encoding CBHIIa

```

ggatccacac ctaccatacc ggatagtatg ctacccaagt gacatagggg ttgtaaagta      60
atacgagaac tcagagagca ctgcccatac ggctcgccaa tgacctcaag tgccagggtca    120
gctttgcgag acagacctga gcgcgtcgga tgtgtgacat ggaacgcgcc ggatcgccctt    180
gttgattaat tatagggaag tagcgaggaa ggtttcagca attgacgtga gcgtacatta    240
aaagctgtat gatttcagga agacgagcca tggaccagggt ttcaaggctg aatggcttga    300
cgacttaagc accgaacgag gaatgaaaga atgaaaagtg ggggatcatt ctggcccctc    360
ctcgtatgtc gagtgttaaa gaaggcgggt ctacggagga cctaaagagc tccaatttgc    420
tctgttgagc ttaagccaca tatctcaaga tgaatacatg tcaggcatag tcaccctgat    480
cttgttcacg agtccacaca cttttcagtt cagcatgttg attcctcatc catatcactt    540
tccattacta tctctttatg tccttgggtc agactccaag gaaccgatag gtgagcatcg    600
gtgaggctcc ctcaaggtac caaagtagcc atcatcaccg aggtctggga atggcgccgt    660
gccgatctg agtccctcaa ctccacggta cgacgacagc acgtcacatt gacgcaccac    720
ggttgaacaa gcagagaggg acacgtcttg ctacgcgaat cctggcactg gatggagacg    780
cgtgtgagca ggtttccgga accatgacgg cctgggtcgg cttctcgaac aaagaagtgg    840
aacacaaaaa gaaccgaaac ggaaacgcag gcacggcatc gacgaccgga ttgtccacg    900
gggacctcgg ccagtcaagc gttgccctgg ccgtcagctc cctggcgacg gggattcagc    960
acatctcacg ttataggcga cctcatcccc ctccgtctt gtgcgggtcg tgcctcgtgc   1020
cgagtacca ggctgtccgg ggccctttagc cggggcgga tcagagtcaa gatgcggccg   1080
aattggacgg cagacgaagt ttctgtaggg gtcgatgatc gcatgacga caccaccccc   1140
tgcgtgatcc cgtggccctg ggctgggaat tgccggctaa taatctacgg cttaatagat   1200
atgcactttg cagcgggtgc agataaataa gctgtggttt caaacactgg cctccgtact   1260
ttaccacca actgccgctt agcgcgggga cctgagctt gggagtgcgc ggagcggcag   1320
ccacctcggg ttagcgtaca cagcaggggt gcatgcgggg atgcccggtg catggcttca   1380
tagtgtacga cagaccgtca agtccaaatc tgggtgatgc ttgatgagat gacagcgagc   1440
cccgctcggc gcaccccggc tatgcatcgc gaattgacaa cactctcagc tctattgcga   1500
cccatcggat aaaagaagaa gaaaaaatg gaccttgagt acgggcgtca gaaacaaaaa   1560
aaaaactcgg gaaccaaata tgtcgggcat ggccgggggtg aacgaccgct actccccgtt   1620
cccttcttcg caaacagaac gctacagagg gttttctggt ttgtcaaaga gttcggaggt   1680
cctctgctcc gcgaatgcgt ggtgaaccca ccagcagcca ttgttcttgc atgcgtggcg   1740
gaccgttagc cgtgatcga catggcgagc ttcccacctc agacctggag cagacgggtg   1800
cgaggagcaa ggggctgccc tccccctgac ggctcgaccc caatgacttc cccaaacggg   1860
gacatcgagg gtcgtgcatg atggtggaaa gtagttgcag tatgggaagt accccgggtt   1920
gccaggaacc gttgttcggc cccccacatt ttctctctgc catgtcaact gtgtgtcgtt   1980
cgagagttcc tggctccggc ccccggtcca attccctaac gggaccgcgg ggcacgcct   2040

```

Serial No. PCT/US2006/027347Docket No. 3123-4008PC

FIGURE 11 (CONT'D)

SEQ ID NOs: 7 and 8 eg6 gene encoding CBHIIa

```

gtaactaact tccaaatgaa gccgatatg agggagggag attggatctg gcaagccagc 2100
cattcgctgc gatcggcact cgtccgtcag ccccgagtc catatcccca aaggcaactg 2160
ctcggcgcg ctcagtctt ctccggaacg tccagcccga aggcgcgcgc cagcaccggc 2220
cctatgttcc tgattgcgat cctcgatctc cagagacggg tcacctcgcc tcgaggacgg 2280
tgcaggggca toggcttcgc ttcctagagc tccgggctgt gtgtggtcaa ggggagaagg 2340
cggcgggccc aaggtgcgtc tcggcgcaact caccatcgcc ctttaccccc cccccccca 2400
gtatataaaa gatggccatc gtctcctcgt ctgcttggga agaaaggatc tctcgaccat 2460
gcaccacagc ctagctctaa cccagcttgt cgtgtgttgt tgcccagc atg aag ttc 2517
                                     Met Lys Phe
                                     1
gtg cag tcc gcc acc ctg gcg ttc gcc gcc acg gcc ctc gct gcg ccc 2565
Val Gln Ser Ala Thr Leu Ala Phe Ala Ala Thr Ala Leu Ala Ala Pro
      5              10              15
tcg cgc acg act ccc cag aag ccc cgc cag gcc tcg gcg ggc tgc gcg 2613
Ser Arg Thr Thr Pro Gln Lys Pro Arg Gln Ala Ser Ala Gly Cys Ala
      20              25              30              35
tcg gcc gtg acg ctc gat gcc agc acc aac gtg ttc cag cag tac acg 2661
Ser Ala Val Thr Leu Asp Ala Ser Thr Asn Val Phe Gln Gln Tyr Thr
              40              45              50
ctg cac ccc aac aac ttc tac cgt gcc gag gtc gag gct gcc gcc gag 2709
Leu His Pro Asn Asn Phe Tyr Arg Ala Glu Val Glu Ala Ala Ala Glu
              55              60              65
gcc atc tcc gac tcg gcg ctg gcc gag aag gcc cgc aag gtc gcc gac 2757
Ala Ile Ser Asp Ser Ala Leu Ala Glu Lys Ala Arg Lys Val Ala Asp
              70              75              80
gtc ggt acc ttc ctg tgg ctc gac acc atc gag aac att ggc cgg ctg 2805
Val Gly Thr Phe Leu Trp Leu Asp Thr Ile Glu Asn Ile Gly Arg Leu
              85              90              95
gag ccc gcg ctc gag gac gtg ccc tgc gag aac atc gtg ggt ctc gtc 2853
Glu Pro Ala Leu Glu Asp Val Pro Cys Glu Asn Ile Val Gly Leu Val
      100              105              110              115
atc tac gac ctc ccg ggc cgt gac tgc gcg gcc aag gcc tcc aac ggc 2901
Ile Tyr Asp Leu Pro Gly Arg Asp Cys Ala Ala Lys Ala Ser Asn Gly
              120              125              130

```

Serial No. PCT/US2006/027347Docket No. 3123-4008PC

FIGURE 11 (CONT'D)

SEQ ID NOs: 7 and 8 eg6 gene encoding CBHIIa

```

gag ctc aag gtc ggc gag ctc gac agg tac aag acc gag tac atc gac a 2950
Glu Leu Lys Val Gly Glu Leu Asp Arg Tyr Lys Thr Glu Tyr Ile Asp
      135              140              145
gtgagttaac cctttgtggc cccttctttt ccccgagag agcgtctggt tgagtggggt 3010
tgtgagagag aaaatggggc gagcttaaag actgacgtgt tggctcgcag ag atc 3065
                                   Lys Ile

gcc gag atc ctc aag gcc cac tcc aac acg gcc ttc gcc ctc gtc atc 3113
Ala Glu Ile Leu Lys Ala His Ser Asn Thr Ala Phe Ala Leu Val Ile
150              155              160              165
gag ccc gac tcg ctc ccc aac ctg gtc acc aat agc gac ctg cag acg 3161
Glu Pro Asp Ser Leu Pro Asn Leu Val Thr Asn Ser Asp Leu Gln Thr
      170              175              180
tgc cag cag agc gct tcc ggc tac cgc gag ggt gtc gcc tat gcc ctc 3209
Cys Gln Gln Ser Ala Ser Gly Tyr Arg Glu Gly Val Ala Tyr Ala Leu
      185              190              195
aag cag ctc aac ctc ccc aac gtg gtc atg tac atc gat gcc ggc cac 3257
Lys Gln Leu Asn Leu Pro Asn Val Val Met Tyr Ile Asp Ala Gly His
      200              205              210
ggg ggc tgg ctc ggc tgg gac gcc aac ctc aag ccc ggc gcc cag gag 3305
Gly Gly Trp Leu Gly Trp Asp Ala Asn Leu Lys Pro Gly Ala Gln Glu
      215              220              225
ctc gcc agc gtc tac aag tct gct ggt tcg ccc tcg caa gtc cgc ggt 3353
Leu Ala Ser Val Tyr Lys Ser Ala Gly Ser Pro Ser Gln Val Arg Gly
230              235              240              245
atc tcc acc aac gtg gct ggt tgg aac gcc tg gtaagacact ctatgtcccc 3405
Ile Ser Thr Asn Val Ala Gly Trp Asn Ala Trp
      250              255
ctcgtcggtc aatggcgagc ggaatggcgt gaaatgcatg gtgctgacct ttgatctttt 3465
ccccctccta tag g gac cag gag ccc ggt gag ttc tcg gac gcc tcg gat 3515
      Asp Gln Glu Pro Gly Glu Phe Ser Asp Ala Ser Asp
      260              265
gcc cag tac aac aag tgc cag aac gag aag atc tac atc aac acc ttt 3563
Ala Gln Tyr Asn Lys Cys Gln Asn Glu Lys Ile Tyr Ile Asn Thr Phe
      270              275              280

```

- 19 of 38 -

Serial No. PCT/US2006/027347Docket No. 3123-4008PC

FIGURE 11 (CONT'D)

SEQ ID NOs: 7 and 8 eg6 gene encoding CBHIIa

```

ggc gct gag ctc aag tct gcc ggc atg ccc aac cac gcc atc atc gac      3611
Gly Ala Glu Leu Lys Ser Ala Gly Met Pro Asn His Ala Ile Ile Asp
285                290                295                300
act ggc cgc aac ggt gtc acc ggt ctc cgc gac gag tgg ggt gac tgg      3659
Thr Gly Arg Asn Gly Val Thr Gly Leu Arg Asp Glu Trp Gly Asp Trp
                305                310                315
tgc aac gtc aac ggc gcc ggc ttc ggt gtg cgc ccg act gcc aac act      3707
Cys Asn Val Asn Gly Ala Gly Phe Gly Val Arg Pro Thr Ala Asn Thr
                320                325                330
ggc gac gag ctc gcc gac gcc ttc gtg tgg gtc aag ccc ggt ggc gag      3755
Gly Asp Glu Leu Ala Asp Ala Phe Val Trp Val Lys Pro Gly Gly Glu
                335                340                345
tcc gac ggc acc agc gac tcg tcg gcg gcg cgc tac gac agc ttc tgc      3803
Ser Asp Gly Thr Ser Asp Ser Ser Ala Ala Arg Tyr Asp Ser Phe Cys
                350                355                360
ggc aag ccc gac gcc ttc aag ccc agc ccc gag gcc ggt acc tgg aac      3851
Gly Lys Pro Asp Ala Phe Lys Pro Ser Pro Glu Ala Gly Thr Trp Asn
365                370                375                380
cag gcc tac ttc gag atg ctc ctc aag aac gcc aac ccg tcc ttc      3896
Gln Ala Tyr Phe Glu Met Leu Leu Lys Asn Ala Asn Pro Ser Phe
                385                390                395
taagctcctc gacggcttct tgctgtcagt cgctctgaag gtggtgtgct ggtggtgccc      3956
ctgctcctgc tgctgtctgt ccgcggggag gggaggcaac gaaaatgaag tcctgcttca      4016
aaacaaaaca gaaacaagcg aggcgcggtg caatggctcg gcgttcgtct tttttcatgt      4076
tcccttctag tgtagtagtt tgatagtcgt acataagggg tttcagaacc gtctctctgt      4136
ctcggtcttt ttgcgagttg ttgcgactcg tgattatggc ctttgttgct cgttgcgcca      4196
gagtagaacc acagcgtgtt ggggtagcag cttgctccgt aggacgtagg gaaacaacct      4256
gagactctgg aattgcagtc agcctgcgtc gccctcttag gaaacgaagg ggagaaccag      4316
tagtggctgc agcttacaaa cgcgagcatg gtgaacatct ccgagaaaag ggagggatcc      4376

```

Serial No. PCT/US2006/027347Docket No. 3123-4008PC

FIGURE 12

SEQ ID NOs: 9 and 10 eg2 gene encoding EGII

```
1      tgctgctctgatgtgctgatgcacagcttcccctcgcgattgccggcaggatctccaacc
61     ctctggatcggagcagacgatcagcgggcacaatggccagcttgccagcggtcaactcca
121    agttgacccgcttttatcacgcccagctggacatgcacaggcttggttctctgtgttcc
181    tacgatctgcacagtaggtttgactgctgatcttcgctttcctgtgcgcctccccctcc
241    ctacgggtaccttatccttgccctgtaaccccgcttatgtcaaacttgagtttgaccaa
301    tgctagcgcaaaagtacctacatagtactatgtaataaggtaggtacatacatcagtagg
361    cgtttatctagtaaattttggctttttgaaactcaattgctcctctcctcgcctccacct
421    ctgcttggaattgacaaccctggctgtgcctagaggtagcatcgacgatcaatcaaatc
481    taaagtattcgagattgacctttctgctctaattatattaattatccgcacaatgctgta
541    gtcatgactctcctttcaagttgccttctcgtttatgtatgtacaatggcggtcatgc
601    ttcatgccaacagatggttctatcggaacaatgtttgactttctggtcgccccgtcgaac
661    tgttttgatttcgcacgggaagtgttcttaccaaagctaagtcgactcgtggagcttcgt
721    aacggccagtgatcgttgatcgcttttgaggaggttgcgatggagcgagaccggctacga
781    gcacgttcgcaaaggcagcagcatagacgacctccgtggcgccattcgggagatgcaca
841    tgacataagcatatcaatactcacctgaactcatcggccgatgcctcgcaggtagttaca
901    agacataatttggtgggtatattatcccaaccgtacctttgtcgcgtcatttcgggtatg
961    tgctgatgcctacttagggagcaaagacgcctctcctcacctgcgggttacttacttact
1021   gtgcagcatggccttatgttctccgggtcttgcttgcggaatgaacaaaaacgcccga
1081   agaaaaagcgcttcttcagattgtgtctaccgaacataagagggttatgtcgagaccg
1141   ccagcaaatgtcaacaaccacccacggcgttcagaaccttcgaaatatcatctagttt
1201   aagtttaaattgacggcccgagtcacagccgagattcccatattggccgataccagcgctc
1261   cctgtttttccaaaggtgtctcgtcaactggcgcatctgcctacaacgagatataatta
1321   ccgttttcttttgcaaaagggcatgcatggatgtatattatgtcctgcagaacgaga
1381   agcaatcatggtgtaggttttgtgcgggtatggagctaataatattgaacggatctctggt
1441   ccgtcctaaatcgttgaaacgctaggcccaggaggacctgctcgacttggcgaacggaga
1501   tttccaggatgaaaggctcggaacatgtccatccgcggccagcctgaacacttttgctcgt
1561   ttccggaccatcgaccacgaaaacagtgcggttgctggcacagtgcgactcacgatgg
1621   cgatggtccagcccgttcccgcccgatgccacttgacgcgaactctccttcattcggc
1681   ggcccgggcgtgtctggcctattagtagcattttgataccggccttggtcgccgcccgg
1741   ttttcttgccgatacgggaatctcggtggtcccaactccacctgggcacgctctggtg
1801   ccaacatggaacttcgggatgccgctccgggcacagtcgaagcgctttaaatacgaactt
1861   accccacaagaatcgaggcgtaacccggaattagggaacacctggacggcgcaacccctgg
1921   accgaaggcctcgctaaccgggttcctggagccgcatgcgcggctgcccgttgcccgc
1981   tcttgagatgacacttcttttcagcggaggatggtcgggcagggaatgatgtattataa
```

Serial No. PCT/US2006/027347Docket No. 3123-4008PC

FIGURE 12 (CONT'D)

SEQ ID NOs: 9 and 10 eg2 gene encoding EGII

```
2041      gaagcgagccgattccgaaggactcgacccccctctctcgccctgtgtccgccagctaatt
2101      acagcactccttctcgacttgaaacgcccagATGAAGTCCTCCATCCTCGCCAGCGTCT
700              M K S S I L A S V

2161      TCGCCACGGGCGCCGTGGCTCAAAGTGGTCCGTGGCAGCAATGTGGTGGCATCGGATGGC
720      F A T G A V A Q S G P W Q Q C G G I G W

2221      AAGGATCGACCGACTGTGTGTCGGGTACCCTGCGTCTACCAGAATGATTGGTACAGCC
740      Q G S T D C V S G Y H C V Y Q N D W Y S

2281      AGTGCCTGCCTGGCGCGCGTCGACAACGCTCCAGACATCTACCACGTCCAGGCCACCG
760      Q C V P G A A S T T L Q T S T T S R P T

2341      CCACCAGCACCGCCCTCCGTGCTCCACCTCGCCTAGCAAGGGCAAGCTCAAGTGGC
780      A T S T A P P S S T T S P S K G K L K W

2401      TCGGCAGCAACGAGTCGGGCGCCGAGTTCGGGGAGGGCAACTACCCGGCCTCTGGGGAA
800      L G S N E S G A E F G E G N Y P G L W G

2461      AGCACTTCATCTTCCGTCGACTTCGGCGATTTCAGgtacgggccaataataatatattat
820      K H F I F P S T S A I Q
2521      tatagcaggcaggagggagcaggagaagaagggaggggaggtggccaacaatcggaaga
2581      agaccgggaggcactgaccgttgattcctttgtgtaatagACGCTCATCAATGATGGATA
861              T L I N D G Y

2641      CAACATCTTCCGGATCGACTTCTCGATGGAGCGTCTGGTGCCCAACCAGTTGACGTCGTC
881      N I F R I D F S M E R L V P N Q L T S S

2701      CTTGACGAGGGCTACCTCCGCAACCTGACCGAGGTGGTCAACTTCGTGACGAACGCGGG
901      F D E G Y L R N L T E V V N F V T N A G

2761      CAAGTACGCCGTCCTGGACCCGCACAACTACGGCCGGTACTACGGCAACGTCATCACGGA
921      K Y A V L D P H N Y G R Y Y G N V I T D

2821      CACGAACGCGTTCGGACCTTCTGGACCAACCTGGCCAAGCAGTTCGCCTCCAACCTCGCT
```

- 22 of 38 -

Serial No. PCT/US2006/027347Docket No. 3123-4008PC

FIGURE 12 (CONT'D)

SEQ ID NOs: 9 and 10 eg2 gene encoding EGII

```
941      T N A F R T F W T N L A K Q F A S N S L
2881    CGTCATCTTCGACACCAACAACGAGTACAACACGATGGACCAGACCCTGGTGCTCAACCT
961      V I F D T N N E Y N T M D Q T L V L N L
2941    CAACCAGGCCGCCATCGACGGCATCCGGGCCCGCGCGACCTCGCAGTACATCTTCGT
981      N Q A A I D G I R A A G A T S Q Y I F V
3001    CGAGGGCAACGCGTGGAGCGGGCCTGGAGCTGGAACACGACCAACACCAACATGGCCGC
1001     E G N A W S G A W S W N T T N T N M A A
3061    CCTGACGGACCCGAGAACAGATCGTGTACGAGATGCACCAGTACCTCGACTCGGACAG
1021     L T D P Q N K I V Y E M H Q Y L D S D S
3121    CTCGGGCACCCACGCCGAGTGCCTCAGCAGCAACATCGGCGCCAGCGCGTCGTCGGAGC
1041     S G T H A E C V S S N I G A Q R V V G A
3181    CACCCAGTGGCTCCGCGCCAACGGCAAGCTCGGCGTCCTCGGCGAGTTCGCCGGCGGGCGC
1061     T Q W L R A N G K L G V L G E F A G G A
3241    CAACGCCGTCTGCCAGCAGGCCGTACCGGCCTCCTCGACCACCTCCAGGACAACAGCGA
1081     N A V C Q Q A V T G L L D H L Q D N S E
3301    GGTCTGGCTGGGTGCCCTCTGGTGGGCCCGGTCCCTGGTGGGGCGACTACATGTACTC
1101     V W L G A L W W A A G P W W G D Y M Y S
3361    GTTCGgtaagtttctcccttgttcttggtttccccccagtaaggagtcaggcaacat
1121     F
3421    gccaagaccggctcggcttcggttcaaggcggttcgttgtaacactgaagagttccaac
3481    ttccaaccctgttcgtgtcctccgatcagcttcgacgggtgaaggggaagggtttgg
3541    gaggtaggtggaggtcaaaaggaggatccccagatctccacaaacggccctgagcca
3601    acaacagcctctggggtcaaaatgggcgccaaccatacggtcattcactcaggacacctg
3661    ctaacgcgtctctttttttgtttccagAGCCTCCTTCGGGCACCGGCTATGTCAACTAC
1221                                     E P P S G T G Y V N Y
3721    AACTCGATCCTAAAGAAGTACTTGCCGTAAGgggcatgcagcaaggtcgagcgagcatta
```

- 23 of 38 -

Serial No. PCT/US2006/027347Docket No. 3123-4008PC

FIGURE 12 (CONT'D)

SEQ ID NOs: 9 and 10 eg2 gene encoding EGII

```
1241      N S I L K K Y L P *
3781      ttcagggccatctgcttgtgtcggcagggcatcacgtcaacccatcgaatcggacagcgga
3841      atgctccgagatgccatacactaagtctgggtgatgacgtgagaatgctggccctggtcgg
3901      ggggttacggccaacaaaaagcaccggacgctgccgcgcccggataccatggtttcatgt
3961      acatattggttctttgctttcttacgggggggggggggggggggggctctgcagcgttgc
4021      tgagcgattcgtttccaagtataactttgtctggaattgaattttgagtgcattgacc
4081      caatcaaccagctcgggtgtgctcacctccggttacccccctcttctccccctgctcggc
4141      ttggctttcctctccgggtgtggagcacggccacggcggtcccaatccatataagatcgat
4201      ggtatactatggtatacactagcttgggaataaactaatccatacgttaactaatggacg
4261      gattatcctaagggtcacgggtcacggttggatataacacctaggatacgggagagctg
4321      atagaaagggatgtactcgtattgtactgtacaatacaaaagtacagatagcacacgaag
4381      tacggtaggtgggtcccgctagtccggaccaacaatagaacatgcgttcttggggacctg
4441      caggaaagaagggggggggggttgccaagacgcccggggttcaaagaaagccccgggccc
4501      ccgatgagatgagacggacgccggcccaaggagaggccgggtggtcgatcctgcaaatgcc
4561      agcaaaaaaaaaatccataccataatccagtcactttcgtcacactcctgtgaaacgagct
4621      ggagggactgctggaaagggttttgagggttaatcactgtatgtggagcatgccgtacctt
4681      ctgtgcttcgttaacagatagagttccagttgaacacacaaagtcttgcctccgctgcca
4741      gacgtgaaaagaagctcctccgggggagctttaggcaactgggagggctctctcccagggt
4801      tcatgggtgtctgctcttcttcaaatttttatgctgccaccccatttgacagaggtgtgca
4861      caccgttgccaggctcttgccatccggcaaaaagcagaaaagtcgacccatcgctaagaa
4921      aggcggtcggaaggggatcggatgctcattgcccgttagcgtctgccattctgacgctg
4981      cccattgttttgtgtcgcattcgtcttcggatgtcggatcaagagtcctggattttttcc
5041      cctgtgcttccagcctaactctgagcgggagctgggtcgggttctgagtgaggttgcttgt
5101      tgggtggagcagcaaccagccaattcactccccgcattttcgcgcccgccaggcatccc
5161      cggcatgcgtttgggcggtaactactccgtactggggtagggtgaaattggttctccgctc
5221      gcaggaggctcgtgctcggtcaggggagaaacaaagtccaactgctccttccctggcaacaa
5281      tgagaggggggttctattgccaaagcttgacgaaaggagcagccacaaaacccaaaagcag
5341      gttaccttactgtacctgagcttgaacgtcgcgtagcattggagctctcgtctaccggcg
5401      gcgtcacactccattggcaggtcaaggcagtcagtgccagcgacccaacaacgtcaatgc
5461      ttgttaccaccagaattacccgggctgcaacactgcagggggccgcccgcgatgttgatca
5521      ccggttgattacttctcggcccgcaaccgggagatgagaagcagaactttgttctccttt
5581      caaaaaggacctgacttgcggggaacgcactgccggcagtgaggaggatgcacgctagtt
5641      atatgtttcccgccatccccagtcgcccgtcgcgctccgtgaggctcagtttggttccc
5701      gtccgcccacaaaacgagcgggtgcataattacatttcgctccatgtaccgtgcaccctcc
5761      ccggttcgagacgtagta
```

- 24 of 38 -

Serial No. PCT/US2006/027347Docket No. 3123-4008PC

FIGURE 13

SEQ ID NOS: 11 and 12 bgl1 gene encoding BGL

```

1      cgggcctccagttccaggagcttggctctgccgacatactgtgtacactaggaattctct
61     tatgcggggtgtgcgcggggaaatgttggggaactcgagttgggtcatgtggacaagacc
121    aatgggagctgacatcattgtgcgaccggttaaaccggaagctacaacaacattctggat
181    tctacactagtgggaagaggttaagtaattgacgacaagcaagaagcattgccatgttctgc
241    gaaggatgcgggtgtttttgcatgagcaggaagctgtggcttttttagtgctcctttgtgc
301    tcgcccggcgcgacagaacactaccgaaacgcaggggactgcgtgcctctggggtcgaatg
361    ccgatccccatcttcacattcccaccatcgtgttctgttaacgaagccggagcggcgga
421    actcgaagctccactacgtatggatacttgggaccgtacggagtgtgttggtacggatgc
481    ctgcacaagtgtgtgcttcctacgaagacgccaacccacataatacacaaaagctgttg
541    taagtcgagttacctcaggcacgttcgggcaactcgggcaacctgacgagatttccccgc
601    cattccgccaagaggccggcgctgcctgattaggcagctcttggaacaatactatgta
661    gaatggaagctccatccatagtcagctccattggcgggtcccagtgatctcgatggctgga
721    tggctgctctgtacggtacatacatagtaagttctcgcttgagagcccaattcgctgca
781    atagcatctttccccgcagtgcgcggcgccctgggtcccgtccacaatgacctgtct
841    tctggagcttctcgacgaacagatcgcccgcttcttctccacaccaatccgaaccagtc
901    gggagcatggctgcggatgcgacgcagccttccttcgcgctgtacaaacagctccgggaa
961    cgtcgactggtatgtacggactacagtaagtacactacgagtgcacatactgacgaatac
1021   cggcctcagaggaacctggcaggaccctacccacacgaaaccacagcgagaaagcgcaa
1081   tggatcagtaactactgcgaagtaaacgtgggtccggggcaaaggatctgagggcggatcg
1141   ctcgtggggctgcgaggcgaggagagcaaaacaagccagtccctcccggaacctggaaaa
1201   tcacttataaacacacgtcaccggcgccgggtgcgcgccatgtgtcacctccaggctcc
1261   tcccgggcgatgatctctgcgggtgccatcaatcatctcgggttcgccgcagctgcttctt
1321   tctgtgcagtgaacgctctcaaactgcaacgacgctgtccgacatgaaggctgctgcgct
1381   ttccctgcctcttcggcagtagcccttgccgttcaggcgccattgaatcgagaaaggatg
1441   gacgggctttcgtcaaagactcgtccccgatcaacttcccctttcatccagaccacccc
1501   aacctcccagtcctgcttcgagcacgatctcttcgggcagcaccacccacatccact
1561   cagattagcggcgacaccgttgactgttgcaatccgcaatcgacATGCAACTTCCAGCCG
                                     M Q L P A
1621   CAGCCCAATGGCTGCTCACGCTTCCCGGAAAGCCTCACTTGCTGACAATCATCGTCAGG
      A A Q W L L T L P A K A S L A D N H R Q
1681   TTCACCAGAAGCCCCCTCGCGAGATCTGAACCTTTTACCCGTCGCCATGGATGAATCCCA
      V H Q K P L A R S E P F Y P S P W M N P
1741   ACGCCGACGGCTGGGCGGAGGCCTATGCCAGGCCAAGTCCTTTGTCTCCCAAATGACTC
      N A D G W A E A Y A Q A K S F V S Q M T

```

Serial No. PCT/US2006/027347Docket No. 3123-4008PC

FIGURE 13 (CONT'D)

SEQ ID NOs: 11 and 12 bgl1 gene encoding BGL

1801 TGCTAGAGAAGGTCAACTTGACCACGGGAGTCGGgtaagttttgtcattttgtccaggta
L L E K V N L T T G V G

C1 Bgl1 236 for

1861 acatgcaaatggttctgctaacaataacttaccgtagCTGGGGGCTGAGCAGTGCCTCG
W G A E Q C V

1921 GCCAAGTGGGCGCGATCCCTCGCCTTGGACTTCGCAGTCTGTGCATGCATGACTCCCTC
G Q V G A I P R L G L R S L C M H D S P

1981 TCGGCATCCGAGGAGCCGACTACAACCTCAGCGTTCCCTCTGGCCAGACCGTTGCTGCTA
L G I R G A D Y N S A F P S G Q T V A A

2041 CCTGGGATCGCGGTCTGATGTACCGTCGCGGTACGCAATGGGCCAGGAGGCCAAAGGCA
T W D R G L M Y R R G Y A M G Q E A K G

2101 AGGGCATCAATGTCTTCTCGGACCACTCGCCGCCCTTGGCCGCATGCCCGAGGGCG
K G I N V L L G P V A G P L G R M P E G

2161 GTCGTAACCTGGGAAGGCTTCGCTCCGGATCCCGTCCTTACCGGCATCGGCATGTCCGAGA
G R N W E G F A P D P V L T G I G M S E

C1Bgl1 682 rev

2221 CGATCAAGGGCATTCAGGATGCTGGCGTCATCGCTTGTGCGAAGCACTTTATTGGAAACG
T I K G I Q D A G V I A C A K H F I G N

2281 AGCAGGgtgagtagtcaaagacgggcccgtctcggacccgcccgttcaagctgctgactct
E Q

2341 gctgcagAGCACTTCAGACAGGTGCCAGAAGCCCAGGGATACGGTTACAACATCAGCGAA
E H F R Q V P E A Q G Y G Y N I S E

2401 ACCCTCTCCTCCAACATTGACGACAAGACCATGCACGAGCTCTACCTTTGGCCGTTTGCC
T L S S N I D D K T M H E L Y L W P F A

2461 GATGCCGTCCGGGCCGGCGTCGGCTCTGTTCATGTGCTCGTACCAGCAGGTCAACAACCTCG
D A V R A G V G S V M C S Y Q Q V N N S

2521 TACGCCTGCCAGAACTCGAAGCTGCTGAACGACCTCCTCAAGAACGAGCTTGGGTTTCAG
Y A C Q N S K L L N D L L K N E L G F Q

2581 GGCTTCGTTCATGAGCGACTGGCAGGCACAGCACACTGGCGCAGCAAGCGCCGTGGCTGGT
G F V M S D W Q A Q H T G A A S A V A G

2641 CTCGATATGTCCATGCCGGGCGACACCCAGTTCAACACTGGCGTCAGTTTCTGGGGCGCC
L D M S M P G D T Q F N T G V S F W G A

2701 AATCTCACCTCGCCCTCCTCAACGGCACAGTCCCTGCCTACCGTCTCGACGACATGGCC
N L T L A V L N G T V P A Y R L D D M A

Serial No. PCT/US2006/027347Docket No. 3123-4008PC

FIGURE 13 (CONT'D)

SEQ ID NOs: 11 and 12 bgl1 gene encoding BGL

```
2761 ATGCGCATCATGGCCGCCCTCTTCAAGGTCACCAAGACCACCCACCTGGAACCCATCAAC
    M R I M A A L F K V T K T T H L E P I N
2821 TTCTCCTTCTGGACCGACGACACTTATGGCCCGATCCACTGGGCGCGCAAGCATGGCTAC
    F S F W T D D T Y G P I H W A A K H G Y
2881 CAGAAGATTAAATCCACGTTGACGTCCGCGCGGACCACGGCAACCTCATCCGGGAGATT
    Q K I N S H V D V R A D H G N L I R E I
2941 GCCGCCAAGGGTACGGTGCTGCTGAAGAATACCGGCTCTTACCCCTGAACAAGCCAAAG
    A A K G T V L L K N T G S L P L N K P K
3001 TTCGTGGCCGTCATCGCGGAGGATGCTGGGTCGAGCCCCAACGGGCCCCAACGGCTGCAGC
    F V A V I G E D A G S S P N G P N G C S
3061 GACCGCGGCTGTAACGAAGGCACGCTCGCCATGGGCTGGGGATCCGGCACAGCCAACAT
    D R G C N E G T L A M G W G S G T A N Y
3121 CCGTACCTCGTTTCCCCGACGCGCGCTCCAGGCCCGGGCCATCCAGGACGGCACGAGG
    P Y L V S P D A A L Q A R A I Q D G T R
3181 TACGAGAGCGTCTGTCCAACACTACGCCGAGGAAAAGACAAAGGCTCTGGTCTCGAGGCC
    Y E S V L S N Y A E E K T K A L V S Q A
3241 AATGCAACCGCCATCGTCTTCTGTCATGCGGACTCAGGCGAGGGCTACATCAACGTGGAC
    N A T A I V F V N A D S G E G Y I N V D
3301 GGTAACGAGGGCGACCGTAAGAACCTGACTCTCTGGAACAACGGTGATACTCTGGTCAAG
    G N E G D R K N L T L W N N G D T L V K
3361 AACGTCTCGAGCTGGTGCAGCAACACCATCGTCTCATCCACTCGGTCGGCCCGGTCCTC
    N V S S W C S N T I V V I H S V G P V L
3421 CTGACCGATTGGTACGACAACCCCAACATCACGGCCATTCTCTGGGCTGGTCTTCCGGGC
    L T D W Y D N P N I T A I L W A G L P G
3481 CAGGAGTCGGGCAACTCCATCACCGACGTGCTTTACGGCAAGGTCAACCCCGCCGCCCGC
    Q E S G N S I T D V L Y G K V N P A A R
3541 TCGCCCTTCACTTGGGGCAAGACCCGCGAAAGCTATGGCGCGGACGTCCTGTACAAGCCG
    S P F T W G K T R E S Y G A D V L Y K P
3601 AATAATGGCAATGGTGCGCCCAACAGGACTTCACCGAGGGCGTCTTCATCGACTACCGC
    N N G N G A P Q Q D F T E G V F I D Y R
3661 TACTTCGACAAGGTTGACGATGACTCGGTTCATCTACGAGTTCGGCCACGGCCTGAGCTAC
    Y F D K V D D D S V I Y E F G H G L S Y
3721 ACCACCTTCGAGTACAGCAACATCCGCGTCGTCAAGTCCAACGTACGCGAGTACCGGCCC
    T T F E Y S N I R V V K S N V S E Y R P
3781 ACGACGGGCACACGGCCCAGGCCCGACGTTTGGCAACTTCTCCACCGACCTCGAGGAC
```

- 27 of 38 -

Serial No. PCT/US2006/027347Docket No. 3123-4008PC

FIGURE 13 (CONT'D)

SEQ ID NOs: 11 and 12 bglI gene encoding BGL

```
      T T G T T A Q A P T F G N F S T D L E D
3841  TATCTCTTCCCAAGGACGAGTTCCCTACATCTACCAGTACATCTACCCGTACCTCAAC
      Y L F P K D E F P Y I Y Q Y I Y P Y L N
3901  ACGACCGACCCCGGAGGGCCTCGGCCGATCCCCACTACGGCCAGACCGCCGAGGAGTTC
      T T D P R R A S A D P H Y G Q T A E E F
3961  CTCCCGCCCCACGCCACCGATGACGACCCCGAGCCGCTCCTCCGGTCCTCGGGCGGAAAC
      L P P H A T D D D P Q P L L R S S G G N
4021  TCCCCCGCGCGCAACCGCCAGCTGTACGACATTGTCTACACAATCACGGCCGACATCACG
      S P G G N R Q L Y D I V Y T I T A D I T
4081  AATACGGGCTCCGTTGTAGGCGAGGAGGTACCGCAGCTCTACGTCTCGCTGGGCGGTCCC
      N T G S V V G E E V P Q L Y V S L G G P
4141  GAGGATCCCAAGGTGCAGCTGCGCGACTTTGACAGGATGCGGATCGAACCCGGCGAGACG
      E D P K V Q L R D F D R M R I E P G E T
4201  AGGCAGTTCACCGGCCGCTGACGCGCAGAGATCTGAGCAACTGGGACGTCACGGTGCAG
      R Q F T G R L T R R D L S N W D V T V Q
4261  GACTGGGTCAATCAGCAGGTATCCCAAGACGGCATATGTTGGGAGGAGCAGCCGGAAGTTG
      D W V I S R Y P K T A Y V G R S S R K L
4321  GATCTCAAGATTGAGCTTCCTTGAATGAGTTTCATCAGGGGCTGCAGAGGGATGGTAACA
      D L K I E L P *
4381  CGTCTTAATCAGAAGTATGATGGAGAAAAGCACTTGGCAAGTTCGGTGAGCAAAAAGA
4441  AGGCACTTATTAAGTGTAGGGCGGTGTTCTATGTTTAAAGGTGCTATGTTTACATATAA
4501  TTAGTATATAATGATTTAATAATTATGTTTAGCAGTTGCTAATGTCTGTAATTTTCGGCGT
4561  GTGATGACTGCTACAACACTGGTCTGTCTTCTAGTCGCCATTGTTAATTATGAAGGTTA
4621  TTGTCTACAATTTCTAATACCTTATGGATGATTGCCCAGCTGGTTTCAAACCTCGTTACGC
4681  GCAAATGGTACGATTGAGGTATTATTCATTGTAAGTACCTCCGTACAGCGTCCCCAACTA
4741  TTTCCATTACAGAGATGCCTCGCTTTTCGGTGCTTTTCGGAACAGGGCTGGCAGCGGATCA
4801  TGGCGCGATCAAAACATGGCGAGCAGCTGTCCAGGACGGAGGACAGGTTGGGGACTGATG
4861  CCTCCCGGACGCATTAAGGTGAGAAGATAGACACGTTTACACAGCGTTGAGACCGACAA
4921  GCCACATTAGGCAGCGCCGGTTGCACCACCGCCGTACGGGCAACGGTTCAATCAATCGA
4981  CAACAGTGGAAAGACAAAGTACTGAAGATCAGGTATTAATAGTGTGAGAGAGAAACAGACG
5041  GTGGAAGTGGGTGCTAATATTTCTCTTGATTTTCGGTGTCCATGGTAGTACAGAACACAA
5101  GAAAAAGAAGGAGGAGTGAGCGGAGAAGGAGGAGGGGGAAGCCAGAAAAAGAACATGAA
5161  AAAGCATACACATTGGAGTCGGTCAGTCGGTTGATTGGTTTGGTAGAGAGCGAAAAAGCA
5221  AGCGTCACCTGTAGGATTCGAACCTACGCTCCCGAAGGAAGTGCCTAAGAACGCTAAGCA
```

Serial No. PCT/US2006/027347Docket No. 3123-4008PC

FIGURE 13 (CONT'D)

SEQ ID NOs: 11 and 12 bgl1 gene encoding BGL

```
5281      AGGTTAGCAGGGCAGCGCGTTAACCACTCCGCCAAAGTGACTGTCGTTGATCATGGTCCA
5341      ATTCAAGTAGCTTATAGGAGTTCAACCAGATCACAAATGCATAGGTGCTCGTAGAACGGT
5401      CTAAGTATGAGTTGATTATAAGCAACCGAATGGCTCTCAGCGGCAACACCGTAGCTGAAG
5461      TAACAAAACGCACCTTTGGTTACTTTCTGACTATAAAAATGGGATATTTGGAAATGACCA
5521      CCCGATAAGGTGTCAAATTCCTAAATGACTGTCTGGGTGTGAAGATGTTACTGTGGTTCCA
5581      CCACGAACCAGTTTTAGTATCCGCATGCTTCAGTCTCTGCGCCTCGACAGGCGGAGGGTG
5641      TGTGTTAGATCAGAATCGATGTGACGCTGTGACCGCGAGGCTCTCGAGCCTAGGTGCGGT
5701      AGTTCTGTTCAAAAAGAAGTGTGTGGCCGGGTTTGGGCGCCCTTATAGCCTACCATCCTG
5761      GCTGTGGTTCCCGAGCGGGAGCCGGTTCTCCGTTTTGGTTCCGATAAAGTGTATATCTG
5821      CCTCCCGGTTTCGCATCTAATTTCTGACTTCGTTCCGGACCTCTGGAGACGTAGGGATAG
5881      GTATGGGATATGCCCGGCATTTTCGTAAATGTCCATAGTCTCTTTCCGGACGAGGCGGCAA
5941      GCTCTCAGAGCTATCTAAGCTTAACCAACCCCTGATCCTTAACCCCTCCAGACCACACCT
6001      CCTGGGAGAATAAACCGGGCTCCAAGATCGAAATCGAAATCAGTGCGCGAACTTGAAATC
```

Serial No. PCT/US2006/027347

Docket No. 3123-4008PC

FIGURE 14

SEQ ID NOs: 13 and 14 eg5 gene encoding EGV

atg cat ctc tcc gcc acc acc ggg ttc ctc gcc ctc ccg gcc ctg gcc	48
Met His Leu Ser Ala Thr Thr Gly Phe Leu Ala Leu Pro Ala Leu Ala	
1 5 10 15	
ctg gcc cag ctc tcg ggc agc ggc cag acg acc ccg tac tgg gac tgc	96
Leu Ala Gln Leu Ser Gly Ser Gly Gln Thr Thr Arg Tyr Trp Asp Cys	
20 25 30	
tgc aag ccg agc tgc gcc tgg ccc ggc aag ggc ccc tcg tct ccg gtg	144
Cys Lys Pro Ser Cys Ala Trp Pro Gly Lys Gly Pro Ser Ser Pro Val	
35 40 45	
cag gcc tgc gac aag aac gac aac ccg ctc aac gac ggc ggc tcc acc	192
Gln Ala Cys Asp Lys Asn Asp Asn Pro Leu Asn Asp Gly Gly Ser Thr	
50 55 60	
cgg tcc ggc tgc gac gcg ggc ggc agc gcc tac atg tgc tcc tcc cag	240
Arg Ser Gly Cys Asp Ala Gly Gly Ser Ala Tyr Met Cys Ser Ser Gln	
65 70 75 80	
agc ccc tgg gcc gtc agc gac gag ctg tcg tac ggc tgg gcg gcc gtc	288
Ser Pro Trp Ala Val Ser Asp Glu Leu Ser Tyr Gly Trp Ala Ala Val	
85 90 95	
aag ctc gcc ggc agc tcc gag tcg cag tgg tgc tgc gcc tgc tac gag	336
Lys Leu Ala Gly Ser Ser Glu Ser Gln Trp Cys Cys Ala Cys Tyr Glu	
100 105 110	
ctg acc ttc acc agc ggg ccg gtc gcg ggc aag aag atg att gtg cag	384
Leu Thr Phe Thr Ser Gly Pro Val Ala Gly Lys Lys Met Ile Val Gln	
115 120 125	

Serial No. PCT/US2006/027347Docket No. 3123-4008PC

FIGURE 14 (CONT'D)

SEQ ID NOs: 13 and 14 eg5 gene encoding EGV

```

gcg acc aac acc ggt ggc gac ctg ggc gac aac cac ttt gac ctg gcc      432
Ala Thr Asn Thr Gly Gly Asp Leu Gly Asp Asn His Phe Asp Leu Ala
      130              135              140

gtgagttgcc tccccctctc cccggaccgc tcagattaga tgagattaga ctttgctcgt      492

aaatcgggtcc aagattccct tgactgacca acaaacatca tacgggcag atc ccc ggt      550
                                   Ile Pro Gly
                                   145

ggc ggt gtc ggt att ttc aac g gtaagctggt gcccccgac ccctccccgg      602
Gly Gly Val Gly Ile Phe Asn
      150

acccctcccc cttttcctcc agcgagccga gttgggatcg ccgagatcga gaactcacac      662

aacttctctc tcgacag cc   tgc acc gac cag tac ggc gct ccc ccg aac      711
              Ala Cys Thr Asp Gln Tyr Gly Ala Pro Pro Asn
              160              165

ggc tgg ggc gac cgc tac ggc ggc atc cat tcc aag gaa gag tgc gaa      759
Gly Trp Gly Asp Arg Tyr Gly Gly Ile His Ser Lys Glu Glu Cys Glu
      170              175              180

tcc ttc ccg gag gcc ctc aag ccc ggc tgc aac tgg cgc ttc gac tg      806
Ser Phe Pro Glu Ala Leu Lys Pro Gly Cys Asn Trp Arg Phe Asp Trp
      185              190              195

gtacgttgct ttgacatacc ggaacccaat tcctccaacc ccccccttt tctccccaa      866

ctccgggggt agtcggaatg tcgcgactga ccctatttca g g ttc caa aac gcc      920
                                   Phe Gln Asn Ala
                                   200

```

Serial No. PCT/US2006/027347Docket No. 3123-4008PC

FIGURE 14 (CONT'D)

SEQ ID NOs: 13 and 14 eg5 gene encoding EGV

gac aac ccg tcg gtc acc ttc cag gag gtg gcc tgc ccg tcg gag ctc	968
Asp Asn Pro Ser Val Thr Phe Gln Glu Val Ala Cys Pro Ser Glu Leu	
205	210
	215
acg tcc aag agc ggc tgc tcc cgt taa	995
Thr Ser Lys Ser Gly Cys Ser Arg	
220	225

Serial No. PCT/US2006/027347Docket No. 3123-4008PC

FIGURE 15

SEQ ID NOs: 15 and 16 eg7 encoding EG VI

```
1      GCGCTTCCGGCCTGGGCGAGTAAATGACGGAAGCCggggccccgtccgactgcgtttgtc
61     ccaactcggaagcaggcatcggtttttgggggggaggaagcgttgcaacacgcactatcg
121    ccaaggtggactcggcgcaatctggaggttcggcccgaggagcgggaatccgggctgaa
181    tctgcgcaaaggctgaccctgcatggtgggaaaatgtaaatatgtgaagttataggcat
241    ataggactcagcgatgacatggaaattgcagaggcatgtgggatttcagcgtttggcatg
301    cattggtcggatctctcgcttctgtctgatgtgatcccgccggagggtgtttcggtctctgg
361    ggaagggacccccctggccccccacctgccccgcacatgcctcgccacgactcccgcg
421    cgccgaggaagaacttcgggtctttgtgacgggagattccactgagtgcgcatcggccaa
481    ccaagcacacaattactccgtacatacacagtacttctgactccgtaaagtaaacgtgt
541    gtttcaaagatcggaatccgtaacagggtactccgtatctaaggtaaatctaccctgtgc
601    acggagcagaacctgaactcttccccctcttactcgagttagtcaccctactccaacca
661    gcggttttcaactcgcaaagtcttgtttataacagtgcataacctgcatttcgtatct
721    cgctagtgtaaagacgaccacacgaggacaaaagaaaaatccaattgcccgatggct
781    cttagtttgaggacagcagcgaaggactacactgcgcgtagtgaccaggccaagaaacg
841    cgaatcgtatattaacggcaaatcaaatggattatagccatttcgcttcggggtgcg
901    tgctcgccgaagtctgggtgccgatcgattgcgaacccccggaatcgcgggatgattcct
961    acagccgcgaaaggggggggggggggggggggtctggacgggacgtgcataaacttcgaa
1021   tttctagaatattgcggattgggttccttcagccctgcgagcgcccccttctggaac
1081   cgcacccttcaccggttcacacacagaggacatgggtggaaatgtgtacctgacgggtg
1141   cccctttgggacagtggagaggcggtatgtcgataaccatccggagccgcagtgtcgac
1201   caagatcttggcttaccatcgacaccaacatgcggactcgtccctcagtcattggagcctt
1261   ggctcgcgagcctccgttcgaagcggctatccgtcctgccagcggaggatctcgtacc
1321   gcttcgcgaactgtgaatgtcctgggtataagagcatggcgcgaccttgtctcgtcagg
1381   aacggggaggaggaggccttgggttagggtcgcgctcgtttggagattgctgagctctgag
1441   ccttcgggtccttggtacccctgcggtccccgggtctcctctctctctctctctctctc
1501   tctctctctctctccacgctcgttcgacagacgcctcccccttcttcgctctccttccc
1561   tcgcacgtagcacactaatagtgcaccATGCGCGTCTCTAGTTTGGTCGCGGCCCTTGCT
      M R V S S L V A A L A

1621   ACCGGTGGTCTTGTGCGCGCCACGCCTAAGCCCAAGGGGTCGTGCCCCCTGGGGCCGTG
      T G G L V A A T P K P K G S S P P G A V

1681   GACGCGAACCCTTTCAAGGGCAAGACGCAGTTCGTCAACCCGGCATGGGCGGCCAAGCTG
      D A N P F K G K T Q F V N P A W A A K L
```

Serial No. PCT/US2006/027347Docket No. 3123-4008PC

FIGURE 15 (CONT'D)

SEQ ID NOs: 15 and 16 eg7 encoding EG VI

1741 GAACAGACCAAAAAGGCGTTCCTGGCCAGGAACGACACCGTCAATGCCGCCAAGACGGAG
 E Q T K K A F L A R N D T V N A A K T E

1801 AAGGTCCAGCAGACCAGCTCGTTCGTCTGGGTCTCGAGGATCGCCGAGCTCTCCAACATC
 K V Q Q T S S F V W V S R I A E L S N I

1861 GACGACGCCATCGCGGCTGCCCCGAAGGCGCAGAAGAAGACGGGCAGGAGGCAGATCGTC
 D D A I A A A R K A Q K K T G R R Q I V

1921 GGCCTGGTGCTCTACAACCTTCCGGACCGCGACTGCAGCGGGCGAGAGCGCGGGCGAG
 G L V L Y N L P D R D C S A G E S A G E

1981 CTCAGCAGCGACAAGAACGGGCTCGAGATCTACAAGACTGAGTTCGTCAAGCCCTTCGCC
 L S S D K N G L E I Y K T E F V K P F A

2041 GACAAGGTGGCGGCCGCAAAGGACCTCGACTTCGCCATCGTCCTGGAGCCCGACTCGCTG
 D K V A A A K D L D F A I V L E P D S L

2101 GCCAACCTGGTCACCAACCTGGGCATCGAGTTCGCGCCAACGCCGCCCCCGTCTACCGC
 A N L V T N L G I E F C A N A A P V Y R

2161 GAGGGCATCGCCTATGCCATCTCCAGCCTTCAGCAGCCAAACGTGCACTTGTACATCGAT
 E G I A Y A I S S L Q Q P N V H L Y I D

2221 GCTGCCACACGGCGGCTGGCTCGGCTGGGACGACAACCTGCCGCTGGCCGCCAAGGAGTTT
 A A H G G W L G W D D N L P L A A K E F

2281 GCCGAGGTGGTCAAGCTTGCCGGCGAGGGCAAGAAGATCCGCGGCTTCGTCACCAACGTG
 A E V V K L A G E G K K I R G F V T N V

2341 TCCAACCTACAACCCCTTCCACGCCGTCGTGCGCGAGAACCTTACCGAGTGGAGCAACTCG
 S N Y N P F H A V V R E N F T E W S N S

2401 TGGGACGAGTCTCACTACGCCTCCTCGCTCACACCGTTCTCGAGAAAGAGGGGCTGCCG
 W D E S H Y A S S L T P F L E K E G L P

- 34 of 38 -

Serial No. PCT/US2006/027347Docket No. 3123-4008PC

FIGURE 15 (CONT'D)

SEQ ID NOs: 15 and 16 eg7 encoding EG VI

2461 GCACGCTTCATCGTCGACCAGGGTCGCGTTGCCCTCCCGGGAGCCCGCAAGGAGTGgtga
A R F I V D Q G R V A L P G A R K E W

2521 gtttcgaccagattgaccctcgacccatgcgaccgagattgctgacgattgaattgcgtg

2581 tcccgccccagGGGTGAATGGTGCAACGTGGCACCCGCCGGATTGGCCCCGCGCCCA
G E W C N V A P A G F G P A P

2641 CGACCAGGGTCAACAACACCGTCGTCGATGCTCTCGTCTGGGTCAAGCCTGGCGGCGAGA
T T R V N N T V V D A L V W V K P G G E

2701 GCGACGGCGAGTGTGGCTTGGCTGGCGCCCCAAGGCCGCCAGTGGTTCGACGAGTACG
S D G E C G L A G A P K A G Q W F D E Y

2761 CCCAGATGCTGGTCGAGAATGCCACCCGCTCTGTCGTCCACAAGTGGTAGataaattttg
A Q M L V E N A H P S V V H K W *

2821 gagtccgagaagggtcccagatagacttttgttttaaaacaaaatgcaagggtgtcgacag
2881 atactggccttaacattaaccaagcaccatgaacatgacttgtcaacatattgatacatc
2941 cgctgctttcccatcgtgctctcaggtctcagggatcaaaggataggtcggtaatgca
3001 aaacgatccattggatccagaagagagaaaaaaaaaggacatgcatgccttgtctgt
3061 catcatgaggaaacaaaggaaaaacaaacgatcgtcgtgttccaacaagctttccaagac
3121 cacaagaccatccaccaacacaacaaacgacaagcaatacgtggaccgccgttgttc
3181 catctctcaagagctgactaaacgaacagtcgttgaaatcatcctacatgagtacgccgc
3241 accacctgttatcgtgtaaaccaaatcgctgttaaagtgcacatctcttaggtatgat
3301 cgtaagttccggtcacggtcacggatcagggatgggtctcaattcgtgtgtcgcgtagcc
3361 gccgccgtatctggacaagacttcttgtattgctccgaaaccgcttttgccgccctaata
3421 atctgtagccttcttacctgggtgggtgccttgaaagacgcggcaggcaacacttcgcaggt
3481 ctgtggcgccaccagcaccaggtgtggtgatgcccggaaaccggtcgtcgacttgctcgc
3541 ggtgtcctcggtcgtggggtgggggtgatgagggttgagggtgttgttgccccgc
3601 aacatccggctccggctccggaccgtccacagacattggacctgcgagcatgactcgtgc
3661 cttcagccagaccaagccatgccatcatcgctctgccgacgctgttgagcgggaggct
3721 gatgttctcagccagaactgcgggtgtacggccatgacctgggctgttcgggtctggcc
3781 gtcttgccggcggtttctccctgccagcttgttgtgcgcgggtgcctgcgagattcgacttc

- 35 of 38 -

Serial No. PCT/US2006/027347Docket No. 3123-4008PC

FIGURE 15 (CONT'D)

SEQ ID NOs: 15 and 16 eg7 encoding EG VI

```
3841      gacctgggcgtggcagaggggtgacgagggacgttgacgccttgatctccttgetcccat
3901      gtccctccaccgtacaggcggaagggtgccatacgcgtccacagcctgcacgagaacct
3961      cagggcgctcgtcaatgagttctgtcaacttgctctcagcctctctatgccgcgagcatc
4021      ctgatcctggagcagaaaccgtgccgagcctccgaggaaacgctccttcagcttcgcgc
4081      gtagtttaggcgtgattcaacaaacgtccggcgaggactcgttggtgcccgacagcagcag
4141      gtccttgatgctgaagccgcgtcggcgaacaggcgcatcatctgggcc
```

Serial No. PCT/US2006/027347Docket No. 3123-4008PC

FIGURE 16

SEQ ID NOs: 17 and 18 xyl2 gene encoding Xyl II

```

1      cgcgccccgtctttgaacgcttgagaagcgacgggtgaagaaccatcaactccgattcc
61     gtcctcatcctcccacgaagccgattgaaatagccacagcggtatgtacggattactc
121    tgctcggtttgacatccatacacagcgctatttttaaaagttcaggacggccaagcccg
181    gttcttggaacggacgacccggattccgaaagctccagcgctcaatgcggtcagtcgtgg
241    cgctgatcctgctgatctgctgatctcataaaccgcgaacttcaacttttcactttgaag
301    cgtatacacgcagcgccctctttcacggcgcatcactcgcaaattaaccgctaatat
361    cctcgcaacttggataaatgtgtagccgacacggaggaggggggttgggggggggttggggg
421    gagacatgatggtctgccccacggatattatttttgttgtttgtataattactgcgg
481    caacattctcaaaggggcccgtgcctcgcggggaaagcccatgacagagaattggacag
541    ctccaagctcgcatatactctaacaacggcgtagctcggaatgaaggcctgccgctcg
601    agtgatagggcgaaagtaaacggacgttacatgcggcacttagccggctgatgccggaga
661    atacgggattcaacgatacaatcacacgatgcgacacacctcggcgacttggcgctctat
721    ggaagaaggctgggttaaagctggcgtagattttgcggtcttggtttcttaaccgggtt
781    atttctattttctcatatgcgcgagcgaatgcggggtgcagagcgccgggagtcgatgg
841    tcctatcagacaagagcctggccccggaacctgggataatagaagccaaattaagccatg
901    ggagtatcgctccgggggtaggaaccgcacgggcaactagaggaggaagaatttgggtataa
961    agggaggacggcggaacaggcttgatggacatgaatcagaagacgacactgggcaactaa
1021   acagcttgacgacagagttttgtgccttgcataggccctcgatatcATGGTCTCGTTCACT
                                     M V S F T
1081   CTCCTCCTCACGGTCATCGCCGCTGCGGTGACGACGGCCAGCCCTCTCGAGGTGGTCAAG
361    L L L T V I A A A V T T A S P L E V V K

1141   CGCGGCATCCAGCCGGGCACGGGCACCCACGAGGGGTACTTCTACTCGTTCTGGACCGAC
381    R G I Q P G T G T H E G Y F Y S F W T D
1201   GGCCGTGGCTCGGTGCACTTCAACCCCGGGCCCCGCGGCTCGTACAGCGTCACCTGGAAC
401    G R G S V D F N P G P R G S Y S V T W N
1261   AACGTCAACAACTGGGTGCGCGCAAGGGCTGGAACCCGGGCCCCGCGCAAGATTGCG
421    N V N N W V G G K G W N P G P P R K I A
1321   TACAACGGCACCTGGAACAACTACAACGTGAACAGCTgtgcggttgctcctcctctttctcc
441    Y N G T W N N Y N V N S
1381   Ctttcgcttggttttccttgatgattgggatccatttttaaagagaaggaaaaaaaaaaca
C1 xyl10 423 for
1441   aaggaaaatagaagataactaacgccaagctctggcagACCTCGCCCTGTACGGCTGGAC
Y L A L Y G W T

```

Serial No. PCT/US2006/027347Docket No. 3123-4008PC

FIGURE 16 (CONT'D)

SEQ ID NOs: 17 and 18 xyl2 gene encoding Xyl II

```

1501    TCGCAACCCGCTGGTCGAGTATTACATCGTGGAGGCATACGGCACGTACAACCCCTCGTC
501      R N P L V E Y Y I V E A Y G T Y N P S S
1561    GGGCACGGCGCGGCTGGGCACCATCGAGGACGACGGCGGCGTGTACGACATCTACAAGAC
521      G T A R L G T I E D D G G V Y D I Y K T
1621    GACGCGGTACAACCAGCCGTCCATCGAGGGGACCTCCACCTTCGACCAGTACTGGTCCGT
541      T R Y N Q P S I E G T S T F D Q Y W S V
1681    CCGCCGCCAGAAGCGCGTCGGCGGCACCTATCGACACGGGCAAGCACTTTGACGAGTGGA
561      R R Q K R V G G T I D T G K H F D E W K
                                C1 Xyl10 722 rev
1741    GCGCCAGGGCAACCTCCAGCTCGGCACCTGGAACTACATGATCGGCCACCGAGGGCTA
581      R Q G N L Q L G T W N Y M I M A T E G Y
1801    CCAGAGCTCTGGTTCCGGCCACTATCGAGGTCCGGGAGGCCTAAagaagccaggcgccctt
601      Q S S G S A T I E V R E A *
1861    cttttgttttgtaggagggggtagaggggggggggagggaaaacgaaaagtagcagggg
1921    ggttttatgccggcagccgtgggccattcgagtgaacctgtatctctctctcccaag
1981    tctccgggctccttctcagagaacttcaatatgtctggggacaaaccaccttgtgaaata
2041    caacggtaattatctaagtttgagtgccctatcgatgcttctgaaaatttcctgctcct
2101    tgatacaagtgcgtttgagccgagccaatgagactgtgtcgattgatagaggccctgaag
2161    gatcaagcgcatgaacaattaagcatgactacgtgcctagctgcagataaatggaagc
2221    cactcaccaaggtcaaccccgcatactggcacgtaagaaccttcctgttacaaggcccaa
2281    ccgactcacatatctatctgcttgggttttgggatgcggtttttaccacaaaaaaaat
2341    ttgatacaatgctctgctgtgcccgggttgctgagaccaagccgtaatcagcgggcaggg
2401    aatcgagtaggtcacgcctgttgccttggtctagaacaaactaatattaaaaagccttg
2461    ctcggcacacatacagaactcgacctgaggcatgttcttgaaggcggctagccagtcaa
2521    gtctggcaccaggccttggtctcgctcagagataccgagggcgaggaggatgaggaagacc
2581    tctttctgcctcagatctcttaggggacgaagaagacaacgccggagccacacaataat
2641    taggtctcatatcagacgtttcggcctggccgagctaataatgtctaattatgcccatcag
2701    ccgtatgtcgaggcaggttgccacgatacgtcgcgcgcgcgcctcattcatctccgact
2761    gggcacaatgtcgccatctcgccgtcaagtggtgcaagatacctattatgcaagcaga
2821    ggatcagatggcgggcccatacagagcggctgctccggcttgcgagaaagccgcttcgcag
2881    caaggtatcgtaggcaggccgcatcttcgggtgggtattctttgtcttgttgccttcgta
2941    attatgtcctggctggcattgtgggaaggggcgaacctcttgatttccgatgggggtcga

```

3123-4008PC.ST25.txt
SEQUENCE LISTING

<110> Dyadic International, Inc.
GUSAKOV, Alexander V.
SALANOVICH, Tatyana N.
ANTONOV, Alexey I.
USTINOV, Boris B.
OKUNEV, Oleg N.
BURLINGAME, Richard P.
EMALFARB, Mark A.
BAEZ, Marco
SINITSYN, Arkady P.

<120> CONSTRUCTION OF HIGHLY EFFICIENT CELLULASE COMPOSITIONS FOR
ENZYMATIC HYDROLYSIS OF CELLULOSE

<130> 3123-4008PC

<140> TBD

<141> 2006-07-13

<160> 18

<170> PatentIn version 3.3

<210> 1

<211> 6360

<212> DNA

<213> Chrysosporium lucknowense

```

<400> 1
ctcagattct aggggtaggg cgggagcaga ggcgaaaatt gggttgtaga atatgaggag      60
ctaggggttg taaactcaaa gaacttcttg ctcttggtct tagtcttctc tcctgggaaa      120
aggggggttt tccgaaagcg gcgctatacg aagccagagg ctactttcct tgctttggat      180
ggcccttgtc caccgttctt gtttcccgtt tgtcaattgc gacgttgccg gcaacctagg      240
tcctaataat taggtagata tttcggtaga ggtagtttaa ttatgcttca gtagagaaat      300
cgttgtctcc acgtctcgca accttgcgaa acttcgccac attgaagata gcattgtctg      360
agttgatttt aaccctttcc agagacgata taatagtgcg agtttctttg atcggaatca      420
tcgacattcg gattttcctt taattatatg aagtattcgg cccacggaac cgggccccga      480
gcagggtgaa ccgcgcaaaa cctcaaccga gtcacctcgc gtccatgttt gtcattggaat      540
caggctccga atcccgtcag atcagtcagt tctggtggct atggacgcgg gagttacggc      600
cagtcgtccc gttgttctgg ggggttgatc aacaggagga agagatctga gatcgaacta      660
caccattga tttatcgacg cataatcaag ttttaataaaa accaaacagc gtgtttggtg      720
ctaccaccga atgcgagatc cgggctagcc cgcggaagga tgatggccac agatctagcg      780
tcatgtatga ttattaccta tgcattctatc ttcgtatctg cctcggggttg gcaaacctg      840
accgagagac gactcgacaa cctgacactt ggcaaaagac atttcgggttg acagcgggag      900
aactccagcg aggaagtcgc ccagagatgc ggatgagaag acaacgccga gacgtgcccg      960

```

3123-4008PC.ST25.txt

cggttggtctt	ccacgaatcg	gagccgactc	ttccgtttgg	ccaatctccg	ggataaatcc	1020
cagcggcggg	tcacgtcacg	tttcatgggg	aggcgcggac	agccatccca	gccaggccat	1080
ggaagagaac	aattcttggg	ggtagcgacc	gagccaaaag	gggggggggg	gaagcgggag	1140
gggaagaagt	ggtattagag	cacgcaccgg	aaaacgcatt	tgggcccttg	ccaacaaaca	1200
ccacaccccc	cgtcctggga	gcaagacatc	caggatgcaa	cccagtaggg	gatgccaaaga	1260
agcatctacg	gcaccatctg	ccggcgcctc	gcctgttaga	gtcccggcac	ccgccaatgg	1320
ggccgtgctg	ggccctgccc	ggcaatgctg	gcgcagcggc	atcaacaaca	ttgctcgggg	1380
agggggcccga	ttttattgat	tagcaaaaaa	acaattaaat	tacccttcca	ttccagcaga	1440
gcttctcctc	cacgcggcgg	cgggaccgct	tgtggacggc	ggtacactac	aaccgcgggg	1500
ctccagtctc	cgtgctgggc	gtgcagatca	cgacccggaa	gagaaatgat	cgcggtctga	1560
cgccgggtac	ggagtactga	gccgcccaacc	acagccgatg	gaccgtgata	tctcaatgcg	1620
ttcaagcaac	acagcaacac	cctggacgag	tctctcctcc	cctaccaccc	cctccccccc	1680
tgccctggcc	gcgaacgggg	cgcgtacccc	agatttctac	tccgtactga	cacccaatc	1740
tattcccgt	ggcgtcgccc	agtctggggc	ggtccggcca	agactctcgg	tgcacgatac	1800
cgcgacgaaa	tcggattaac	cgttggctga	tcaattccaa	gtcaaggagg	aagtggatatg	1860
gaaagtcggc	tcagttttcc	actgcccccg	acaggcagg	tccggatctg	gacagcagtc	1920
ttccgaatct	ttggcagaga	ctcatgataa	tataaaaagg	caaagtaggc	ggcgccttgg	1980
acaggtccat	tctcccaccg	ctcaaccagc	ctccaattcc	tcagaagtct	gttgctctct	2040
cgcagtcgca	gtcaagatga	agcagtacct	ccagtacctc	gcggcgaccc	tgcccctgg	2100
gggcctggcc	acggcccagc	aggcgggtaa	cctgcagacc	gagactcacc	ccaagctcac	2160
ttggtcgaag	tgcacggccc	cgggatcctg	ccaacagg	aacggcgagg	tcgtcatcga	2220
ctccaactgg	cgtgggtgc	acgacgagaa	cgcgcagaac	tgctacgacg	gcaaccagtg	2280
gaccaacgct	tgcagctctg	ccaccgactg	cgccgagaat	tgcgcgctcg	agggtgccga	2340
ctaccagggc	acctatggcg	cctcgaccag	cggcaatgcc	ctgacgctca	ccttcgtcac	2400
taagcacgag	tacggcacca	acattgggtc	gcgcctctac	ctcatgaacg	gcgcgaacaa	2460
gtaccagatg	ttcacctca	agggcaacga	gctggccttc	gacgtcgacc	tctcggccgt	2520
cgagtgcggc	ctcaacagcg	ccctctactt	cgtggccatg	gaggaggatg	gcggtgtgtc	2580
gagctacccg	accaacacgg	ccggtgctaa	gttcggcact	ggggtaagtt	caacgacccg	2640
agacgggtgc	ccttattatc	tgctgcgaaa	acggacggtc	cccttttgct	aactaccctc	2700
ctccaaacag	tactgcgacg	cccaatgcgc	acgcgacctc	aagttcgtcg	gcggcaaggg	2760
caacatcgag	ggctggaagc	cgtccaccaa	cgatgccaat	gccggtgtcg	gtccttatgg	2820
cgggtgctgc	gctgagatcg	acgtctggta	agttttgttg	cctgggcagc	aatggtatat	2880

3123-4008PC.ST25.txt

tagctcgagt ggttcccgtc gttgctgacc ctctcttacc agggagtcga acaagtatgc	2940
tttcgctttc accccgcacg gttgcgagaa ccctaaatac cacgtctgcg agaccaccaa	3000
ctgcggtggc acctactccg aggaccgctt cgctggtgac tgcgatgcca acggctgcga	3060
ctacaacccc taccgcatgg gcaaccagga cttctacggt cccggcttga cggtcgatac	3120
cagcaagaag ttcacgtgag tacaccgtgc ttgaagcccc ctcccccccc ccccccaaaa	3180
aaaaaaagaa aaaagaagtc aaatgattga tgctaaccaa atcaaataac agcgtcgtca	3240
gccagttcga ggagaacaag ctcacccagt tcttcgtcca ggacggcaag aagattgaga	3300
tccccggccc caaggtcgag ggcacgatg cgacagcgc cgctatcacc cctgagctgt	3360
gcagtgccct gttcaaggcc ttcgatgacc gtgaccgctt ctcggagggtt ggcggttcg	3420
atgccatcaa cacggccctc agcactccca tggtcctcgt catgtccatc tgggatgatg	3480
tacgttacct aacccccccc cctttttttt ttcccgcttc tctccccgaa actgccacta	3540
cttatatacg tcccgcgctc atgatgctta ctttttctcc ttccagcact acgccaatat	3600
gctctggctc gactcgagct acccccctga gaaggctggc cagcctggcg gtgaccgtgg	3660
cccgtgtcct caggactctg gcgtcccggc cgacgttgag gctcagtacc ctaatgcgtg	3720
agtcgaaacc gtaaaatgtc gggcaaaaaa aagatcgctc aagctaacga aataatatga	3780
ttagcaagggt catctggtcc aacatccgct tcggcccat cggtctgact gtcaacgtct	3840
aaactgcaac ctgaccgggc cttttctctc cacccccacc cctctcaagt tctctctggt	3900
ggagccctcg tgtccttctt ttcctagggt cgcgaaacct tgagcttggt tatcgtaggg	3960
tcattgtgta catacacaaa aacttaacat ctgctaccaa gatcttggcg ctttgccagg	4020
tctttctcaa cctcgaagca ctgagccttt gtcctccgag tgaagtagga tgactattta	4080
cgttgcaaga ctacgcggtg aaggggacgg agcagacctg ccacagatat tcgtttggtt	4140
gcttgattta tagcagagtc cgaacgtaga catggcccct gaagggtgcca accctagata	4200
gccagaagcc ttgttttacg aaagggtggt caaccaacgg tgctcctcgc tcagcgaatc	4260
taccgcacg caatgtatcg taagaatgtg aactaaaggg aacgacgagg catagggaaa	4320
cgtcaatgtg gcttgaataa cagagttaaa tacctaatag aagaaattag catgccaaga	4380
ttgagccagc aacacatggt agaatagcca gcaaaggacg cttgttcgct tgatctcgaa	4440
ccgtccaacc tgattcgaag gaggaggga aagttgaaga ataccggcaa taattactcg	4500
aggttcctat gccctgcaga gtctaattaa tattaaaggc accaccgcga tgattccgca	4560
attataagca taataagctc gcgggccccca cacgtgcctt caccctccca tgtgtataca	4620
atctgtacct cgttattgtc gaatcgctat tccgatagcg aaggctctggc actcatcaga	4680
taccgtgaca tcgattgaga tttggccggg ccaccggtag taagcgatga gttgggtcatc	4740

3123-4008PC.ST25.txt

aattatcaac aatgcgctca atcagcgata atcagcctat caaccgcgaa atcatacgcg 4800
catcaacgaa ttgtccatca tgcacgtagc ttgtcggcag tgccgcatac cctccagagc 4860
atcatagccg ggatagaaag ctgcgtttca gccgtcccag agtccgagat gcaggtagca 4920
agccttcaag accagttata tgtgacccgg gtaaaatact tggtagagatg caatgggctg 4980
agcttcgggc acttataagc ttacttagat attatctcaa ggtttctttt tgaactcctc 5040
ctagacattt actataaact accgagcttc aatgctagac gccctccttc tgttaaatag 5100
tcttttcctt ctaagagcat ctgccttttt tcccttaggc ttagaggata gggcccctcc 5160
atcttgctgc gacggcctta gccttgggga gtaattattg gtatccgcgt acctgtttcc 5220
cagacagccg aagtttcgac gacaaagtaa ttattgcgac aataccaccg ccatatgcta 5280
ttccgagtgg gtgagccccg aaaacatcgc ttaccgcacg gccatcccag acgacagagg 5340
gcgactttga tgtcttgctc cagatcgccg cacctaacac ggtgggatgg gctggtatcg 5400
tatgggacgg catcatggtc aacaaccccc tgacgggggtg ttgggccaat ggaaacacca 5460
ccgttgctc gagccggatc gcaaggtaag ccgaagagga caaatgacga tgagactttc 5520
tttctttttt attttatttt tttttaaatt tcttttttaa gcgtaatgaa aagagctaca 5580
tatctgtggg tcgttcctca atttcagcga cctctccacc gaagcatcgt caaataagaa 5640
gttgctcgaa acaaagggtg tcagaagcta tagagcttct aaggatatta gccacataca 5700
tgccatagct gtataaggct atttaacgct ttggccagct cctttgtcta taaatattag 5760
tcgttttgct tcctttgtag ataattttaa caaggcactc ttttccttta tatagccacc 5820
tactatagac tgctttcaac gctcccggaa gcttattact acgttcggca gttataagcc 5880
tggcgccttg actactcctc tgccgacgta tctttaatat tagtagtagc ttcttctatt 5940
acgaactctc ttaccctgct ttaatacgct ttcgacgacg tgtctattat atctaagatc 6000
ctagtcgaga cttctatatg ccttactagg cctagtctt agaacttgta gtatattaaa 6060
ctatagttat aggctaaatt tgctagtata tagagatttg ttaaccttaa tagtaattat 6120
aaactagatc tagaagtttt atagtgccta acctataaat aagctagaga taaccttatt 6180
ttagcttcct aggagtaatt cctagaagga gtattacctt taatatctat agatttgata 6240
ccttctaata tagctatcat agctaaattt atataattat aagattcctt ttataaaaat 6300
attatatata ctatagatat tagtaagtag ataggatagc tataatacta gctagtatat 6360

<210> 2
<211> 450
<212> PRT
<213> Chrysosporium lucknowense

<400> 2

Met Lys Gln Tyr Leu Gln Tyr Leu Ala Ala Thr Leu Pro Leu Val Gly

3123-4008PC.ST25.txt

1 5 10 15
 Leu Ala Thr Ala Gln Gln Ala Gly Asn Leu Gln Thr Glu Thr His Pro
 20 25 30
 Lys Leu Thr Trp Ser Lys Cys Thr Ala Pro Gly Ser Cys Gln Gln Val
 35 40 45
 Asn Gly Glu Val Val Ile Asp Ser Asn Trp Arg Trp Val His Asp Glu
 50 55 60
 Asn Ala Gln Asn Cys Tyr Asp Gly Asn Gln Trp Thr Asn Ala Cys Ser
 65 70 75 80
 Ser Ala Thr Asp Cys Ala Glu Asn Cys Ala Leu Glu Gly Ala Asp Tyr
 85 90 95
 Gln Gly Thr Tyr Gly Ala Ser Thr Ser Gly Asn Ala Leu Thr Leu Thr
 100 105 110
 Phe Val Thr Lys His Glu Tyr Gly Thr Asn Ile Gly Ser Arg Leu Tyr
 115 120 125
 Leu Met Asn Gly Ala Asn Lys Tyr Gln Met Phe Thr Leu Lys Gly Asn
 130 135 140
 Glu Leu Ala Phe Asp Val Asp Leu Ser Ala Val Glu Cys Gly Leu Asn
 145 150 155 160
 Ser Ala Leu Tyr Phe Val Ala Met Glu Glu Asp Gly Gly Val Ser Ser
 165 170 175
 Tyr Pro Thr Asn Thr Ala Gly Ala Lys Phe Gly Thr Gly Tyr Cys Asp
 180 185 190
 Ala Gln Cys Ala Arg Asp Leu Lys Phe Val Gly Gly Lys Gly Asn Ile
 195 200 205
 Glu Gly Trp Lys Pro Ser Thr Asn Asp Ala Asn Ala Gly Val Gly Pro
 210 215 220
 Tyr Gly Gly Cys Cys Ala Glu Ile Asp Val Trp Glu Ser Asn Lys Tyr
 225 230 235 240
 Ala Phe Ala Phe Thr Pro His Gly Cys Glu Asn Pro Lys Tyr His Val
 245 250 255

3123-4008PC.ST25.txt
 Cys Glu Thr Thr Asn Cys Gly Gly Thr Tyr Ser Glu Asp Arg Phe Ala
 260 265 270
 Gly Asp Cys Asp Ala Asn Gly Cys Asp Tyr Asn Pro Tyr Arg Met Gly
 275 280 285
 Asn Gln Asp Phe Tyr Gly Pro Gly Leu Thr Val Asp Thr Ser Lys Lys
 290 295 300
 Phe Thr Val Val Ser Gln Phe Glu Glu Asn Lys Leu Thr Gln Phe Phe
 305 310 315 320
 Val Gln Asp Gly Lys Lys Ile Glu Ile Pro Gly Pro Lys Val Glu Gly
 325 330 335
 Ile Asp Ala Asp Ser Ala Ala Ile Thr Pro Glu Leu Cys Ser Ala Leu
 340 345 350
 Phe Lys Ala Phe Asp Asp Arg Asp Arg Phe Ser Glu Val Gly Gly Phe
 355 360 365
 Asp Ala Ile Asn Thr Ala Leu Ser Thr Pro Met Val Leu Val Met Ser
 370 375 380
 Ile Trp Asp Asp His Tyr Ala Asn Met Leu Trp Leu Asp Ser Ser Tyr
 385 390 395 400
 Pro Pro Glu Lys Ala Gly Gln Pro Gly Gly Asp Arg Gly Pro Cys Pro
 405 410 415
 Gln Asp Ser Gly Val Pro Ala Asp Val Glu Ala Gln Tyr Pro Asn Ala
 420 425 430
 Lys Val Ile Trp Ser Asn Ile Arg Phe Gly Pro Ile Gly Ser Thr Val
 435 440 445
 Asn Val
 450

<210> 3
 <211> 3900
 <212> DNA
 <213> Chrysosporium lucknowense

<400> 3
 ccgcaagtga atatgtaatt actcaatgga agttctcgaa acggagtcca gaaatgatgt 60
 ggttctgtgg gaatgcggca agaggcgacg ttgccgtgaa tgcgtgaaca ttcccgctc 120
 ttcttcttct cgtcttcttc cttcttcttc tttcgggtcg cggatggttg acggccagcg 180
 Page 6

3123-4008PC.ST25.txt

tgcgcacggc	tgcggttat	cgagcgtcgg	tacgtctagc	caacatcccc	tagacacgac	240
gaccaagcgt	cttgagaatg	caacaacgtc	tcggaacctg	gcacgcatct	tccgccgcag	300
gtcggcagac	gccgcctggg	caataccacc	cctgtccagg	ccctttcccc	gcaggcagag	360
ccgcgctctt	cctttcatgg	ttattcagga	acgtggcttc	cgagattctc	gcctgttctc	420
ccccagtcaa	cctgccgacc	gtaacccggg	tccaccaccg	cggactgtcc	gcaaaacctg	480
gttcgccccg	gattaatatg	ctatttccgg	actaagtgca	caacacacaa	gcaccccttc	540
cgcctcgcgc	tctagaatct	gctttctaac	ccggttctcg	ggcccttccc	tttcgcgacg	600
cctccgctct	ccttaccagg	caccatccgc	aataggttaag	gtagccaacc	gttttgagac	660
gtgattctgc	caaggaccgc	atccttgcat	tcgccatctg	gtcaaggacc	cctctttccc	720
gctccattct	ggtggctcta	tcgggacggc	gttcccatg	gctctccagg	agagtgatgt	780
gcgagtctgg	agagccgggg	ttggcgtcac	gatgctgccc	acctagggcc	ggccagcccc	840
gcactgcgct	cccgttgatc	cgtctatccc	cgtcaagagc	accagccccg	gcgctcgtga	900
attttcgact	tgttcgactt	gctacagggt	ataaagagga	tgacgcccgc	cctcgatcgg	960
cctgtgtggt	ttctctccct	cgtgccaaac	cactcccacc	tcccgccccg	agatagttgc	1020
ttgtttcgct	ccgtgagagg	gacacacacc	aatggccaag	aagcttttca	tcaccgccgc	1080
gcttgcggt	gccgtgttg	cggcccccg	cattgaggag	cgccagaact	gcggcgctgt	1140
gtggtaagaa	agccccgtcc	gagtctccca	tgattttctc	gtcgagtaat	ggcataaggg	1200
ccaccccttc	gactgaccgt	gagaatcgat	caaatccagg	actcaatgcg	gcggtaacgg	1260
gtggcaagg	cccacatgct	gcgcctcggg	ctcgacctgc	gttgcgcaga	acgagtggta	1320
ctctcagtgc	ctgcccaaca	gccagggtgac	gagttccacc	actccgtcgt	cgacttccac	1380
ctcgcagcgc	agcaccagca	cctccagcag	caccaccagg	agcggcagct	cctcctcctc	1440
ctccaccacg	ccccgccccg	tctccagccc	cgtgaccagc	attcccggcg	gtgcgacctc	1500
cacggcgagc	tactctggca	acccttctc	gggcgctccg	ctcttcgcca	acgactacta	1560
cagggtccgag	gtccacaatc	tcgccattcc	tagcatgact	ggtactctgg	cggccaaggc	1620
ttccgccgtc	gccgaagtcc	ctagcttcca	gtggctcgac	cggaacgtca	ccatcgacac	1680
cctgatggtc	cagactctgt	cccagggtccg	ggctctcaat	aaggccgggtg	ccaatcctcc	1740
ctatgctggg	gagttacatg	gcgacttgcc	ttctcgtccc	ctacctttct	tgacgggatc	1800
ggttacctga	cctggaggca	aaacaacaac	agcccaactc	gtcgtctacg	acctccccga	1860
ccgtgactgt	gccgccgctg	cgtccaacgg	cgagttttcg	attgcaaacg	gcggcgccgc	1920
caactacagg	agctacatcg	acgctatccg	caagcacatc	attgagtact	cggacatccg	1980
gatcatcctg	gttatcgagc	ccgactcgat	ggccaacatg	gtgaccaaca	tgaacgtggc	2040

3123-4008PC.ST25.txt

caagtgcagc aacgccgcgt cgacgtacca cgagttgacc gtgtacgcgc tcaagcagct	2100
gaacctgccc aacgtcgcca tgtatctcga cgccggccac gccggctggc tcggctggcc	2160
cgccaacatc cagcccgccg ccgagctgtt tgccggcatc tacaatgatg ccggcaagcc	2220
ggctgccgtc cgcggcctgg cactaacgt cgccaactac aacgcctgga gcatcgcttc	2280
ggccccgtcg tacacgtcgc ctaaccctaa ctacgacgag aagcactaca tcgaggcctt	2340
cagcccgtc ttgaactcgg ccggcttccc cgcacgcttc attgtcgaca ctggccgcaa	2400
cggcaaacia cctaccggtg tgtttttttt tcttttgtct ctgtcccccc cttttctccc	2460
ccttcagttg gcgtccacaa ggtctcttag tcctgcttca tctgtgacca acctcccccc	2520
ccccggcacc gccacaacc gtttgactct atactcttgg gaatgggagc cgaaactgac	2580
cgttccacag gccacaaca gtgggggtgac tgggtgcaatg tcaagggcac cggctttggc	2640
gtgcgcccga cggccaacac gggccacgag ctgggtcgatg ctttgtctg ggtcaagccc	2700
ggcggcgagt ccgacggcac aagcgacacc agcgccgccc gctacgacta ccactgcggc	2760
ctgtccgatg ccctgcagcc tgcccccgag gctggacagt ggttccaggc ctacttcgag	2820
cagctgctca ccaacgccaa cccgcccttc taaacctcgt cataaagaga gagagatggc	2880
gggcatgggc ctgattgggt tcattgacca tgcggctctt ctgggggtac atattttacc	2940
tacctaccta taaataaggc ggcctatcgg gctctcgctt cgtttattag gtacttggtc	3000
ttgtacatac tttgtttata catacagcag ttagcatcca ctattcgttt cgacaaagcg	3060
gaactttcca gaaaaaaaaa ggttgtacat aattagtctt taggcttcga ttctttgtgc	3120
ctttcttttt ggtaaaaaaa aaattttttt tgaggcatga ttaccttagg tacgttcgtc	3180
gttgtattgg tccccctgca ttttggcgag agagcagctc agccccctgc aaatccctca	3240
acgggagctt aattccctcc actcgggtct tcagcgagac cagccgtcca gagtatcca	3300
gcgtgtagtt gcccacgaa ccagtcgtcc tcgtaagcct cgtcaaagtg tccaagagca	3360
gtatagaagc aacgacctcc gtcaaaagtc tggcaccatg cgatcgggtg gtcctccccg	3420
tgcgccccgc cctcgtagga cttctcatcc acgccaagga gcacgtgcag gccgtcggac	3480
gtcgccccgc ggtgcgcctt gaagttgtac cattcgctct tccagacgag ctccagctgc	3540
gcctgcttgg gttcctgcgg ttcttgcggt tcctgcgctg gccggtcggc gccgccgtct	3600
tggtcacacg cccgcagcga catgactggg tgtttcgggt cgagcagctt gacgagcccg	3660
acctgggggt ccgggtgggt gtcgaacacg gcgccaatga ggtggccgta ccattcggat	3720
gactgcatgg cgaagctggc gcagtgtacc gccacgatcc cgccgccccg ctggacgaaa	3780
ccccgcaggg cgcccagctg cgcgccgtcc aggaactcgc ccgagcactg caggaggacg	3840
atgacgcgat acgccgagag ggagccgggg ctgaacacgg cgggatcctc gctgtcgtcc	3900

3123-4008PC.ST25.txt

<210> 4
 <211> 481
 <212> PRT
 <213> Chrysosporium lucknowense

<400> 4

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

Ala Ala Pro Val Ile Glu Glu Arg Gln Asn Cys Gly Ala Val Thr Gln
 20 25 30

Cys Gly Gly Asn Gly Trp Gln Gly Pro Thr Cys Cys Ala Ser Gly Ser
 35 40 45

Thr Cys Val Ala Gln Asn Glu Trp Tyr Ser Gln Cys Leu Pro Asn Ser
 50 55 60

Gln Val Thr Ser Ser Thr Thr Pro Ser Ser Thr Ser Thr Ser Gln Arg
 65 70 75 80

Ser Thr Ser Thr Ser Ser Ser Thr Thr Arg Ser Gly Ser Ser Ser Ser
 85 90 95

Ser Ser Thr Thr Pro Pro Pro Val Ser Ser Pro Val Thr Ser Ile Pro
 100 105 110

Gly Gly Ala Thr Ser Thr Ala Ser Tyr Ser Gly Asn Pro Phe Ser Gly
 115 120 125

Val Arg Leu Phe Ala Asn Asp Tyr Tyr Arg Ser Glu Val His Asn Leu
 130 135 140

Ala Ile Pro Ser Met Thr Gly Thr Leu Ala Ala Lys Ala Ser Ala Val
 145 150 155 160

Ala Glu Val Pro Ser Phe Gln Trp Leu Asp Arg Asn Val Thr Ile Asp
 165 170 175

Thr Leu Met Val Gln Thr Leu Ser Gln Val Arg Ala Leu Asn Lys Ala
 180 185 190

Gly Ala Asn Pro Pro Tyr Ala Ala Gln Leu Val Val Tyr Asp Leu Pro
 195 200 205

Asp Arg Asp Cys Ala Ala Ala Ala Ser Asn Gly Glu Phe Ser Ile Ala
 210 215 220

3123-4008PC.ST25.txt

Asn Gly Gly Ala Ala Asn Tyr Arg Ser Tyr Ile Asp Ala Ile Arg Lys
 225 230 235 240
 His Ile Ile Glu Tyr Ser Asp Ile Arg Ile Ile Leu Val Ile Glu Pro
 245 250 255
 Asp Ser Met Ala Asn Met Val Thr Asn Met Asn Val Ala Lys Cys Ser
 260 265 270
 Asn Ala Ala Ser Thr Tyr His Glu Leu Thr Val Tyr Ala Leu Lys Gln
 275 280 285
 Leu Asn Leu Pro Asn Val Ala Met Tyr Leu Asp Ala Gly His Ala Gly
 290 295 300
 Trp Leu Gly Trp Pro Ala Asn Ile Gln Pro Ala Ala Glu Leu Phe Ala
 305 310 315 320
 Gly Ile Tyr Asn Asp Ala Gly Lys Pro Ala Ala Val Arg Gly Leu Ala
 325 330 335
 Thr Asn Val Ala Asn Tyr Asn Ala Trp Ser Ile Ala Ser Ala Pro Ser
 340 345 350
 Tyr Thr Ser Pro Asn Pro Asn Tyr Asp Glu Lys His Tyr Ile Glu Ala
 355 360 365
 Phe Ser Pro Leu Leu Asn Ser Ala Gly Phe Pro Ala Arg Phe Ile Val
 370 375 380
 Asp Thr Gly Arg Asn Gly Lys Gln Pro Thr Gly Gln Gln Gln Trp Gly
 385 390 395 400
 Asp Trp Cys Asn Val Lys Gly Thr Gly Phe Gly Val Arg Pro Thr Ala
 405 410 415
 Asn Thr Gly His Glu Leu Val Asp Ala Phe Val Trp Val Lys Pro Gly
 420 425 430
 Gly Glu Ser Asp Gly Thr Ser Asp Thr Ser Ala Ala Arg Tyr Asp Tyr
 435 440 445
 His Cys Gly Leu Ser Asp Ala Leu Gln Pro Ala Pro Glu Ala Gly Gln
 450 455 460
 Trp Phe Gln Ala Tyr Phe Glu Gln Leu Leu Thr Asn Ala Asn Pro Pro
 465 470 475 480

3123-4008PC.ST25.txt

Phe

<210> 5
<211> 1648
<212> DNA
<213> Chyrsosporium lucknowense

<220>
<221> misc_feature
<222> (812)..(812)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (1162)..(1162)
<223> n is a, c, g, or t

<400> 5
atgtacgcc agttcgcgac cctcgccgcc cttgtggctg gcgccgctgc tcagaacgcc 60
tgcactctga ccgctgagaa ccacccctcg ctgacgtggt ccaagtgcac gtctggcggc 120
agctgcacca gcgtccaggg ttccatcacc atcgacgcc actggcggtg gactcaccgg 180
accgatagcg ccaccaactg ctacgagggc aacaagtggg atacttcgta ctgcagcgat 240
ggtccttctt gcgcctccaa gtgctgcatc gacggcgctg actactcgag cacctatggc 300
atcaccacga gcggtaactc cctgaacctc aagttcgtca ccaagggcca gtactcgacc 360
aacatcgggt cgcgctaccta cctgatggag agcgacacca agtaccagag taagtctctc 420
tcgcacccgg ccgccgggag atgatggcgc ccagcccgtg gacgcgaatg acacagtgtt 480
ccagctcttc ggcaacgagt tcaccttcga tgtcgacgtc tccaacctcg gctgcggcct 540
caatggcgcc ctctacttcg tgtccatgga tgccgatggt ggcatgtcca agtactcggg 600
caacaaggca ggtgccaagt acgggtaccg ctactgtgat tctcagtgcc cccgcgacct 660
caagttcatc aacggcgagg ccaacgtaga gaactggcag agctcgacca acgatgccaa 720
cgccggcacg ggcaagtacg gcagctgctg ctccgagatg gacgtctggg aggccaacaa 780
catggccgcc gccttctact cccacccttg cnccgtgatc ggccagtcgc gctgcgaggg 840
cgactcgtgc ggcggtacct acagcaccga ccgctatgcc ggcattctgcg accccgacgg 900
atgcgacttc aactcgtacc gccagggcaa caagaccttc tacggcaagg gcatgacggt 960
cgacacgacc aagaagatca cggtcgtcac ccagttcctc aagaactcgg ccggcgagct 1020
ctccgagatc aagcggttct acgtccagaa cggcaagggtc atccccaaact ccgagtccac 1080
catcccgggc gtcgagggca actccatcac ccaggactgg tgcgaccgcc agaaggccgc 1140
cttcggcgac gtgaccgact tncaggacaa gggcgggcatg gtccagatgg gcaaggccct 1200
cgcgggggccc atggtcctcg tcatgtccat ctgggacgac cacgccgtca acatgctctg 1260

3123-4008PC.ST25.txt

```

gctcgactcc acctggccca tcgacggcgc cggcaagccg ggcgccgagc gcggtgcctg 1320
ccccaccacc tcgggcgtcc ccgctgaggt cgaggccgag gcccccaact ccaacgtcat 1380
ctttctccaac atccgcttcg gccccatcgg ctccaccgtc tccggcctgc ccgacggcgg 1440
cagcggcaac cccaaccgc cgcgcagctc gtccaccccg gtcccctcct cgtccaccac 1500
atcctccggt tcctccggcc cgactggcgg cacgggtgtc gctaagcact atgagcaatg 1560
cggaggaatc gggttcactg gccctaccca gtgcgagagc ccctacactt gcaccaagct 1620
gaatgactgg tactcgcagt gcctgtaa 1648

```

<210> 6
 <211> 526
 <212> PRT
 <213> Chrysosporium lucknowense

<220>
 <221> misc_feature
 <222> (249)..(249)
 <223> Xaa can be any naturally occurring amino acid

<220>
 <221> misc_feature
 <222> (365)..(365)
 <223> Xaa can be any naturally occurring amino acid

<400> 6

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

Ala Gln Asn Ala Cys Thr Leu Thr Ala Glu Asn His Pro Ser Leu Thr
20 25 30

Trp Ser Lys Cys Thr Ser Gly Gly Ser Cys Thr Ser Val Gln Gly Ser
35 40 45

Ile Thr Ile Asp Ala Asn Trp Arg Trp Thr His Arg Thr Asp Ser Ala
50 55 60

Thr Asn Cys Tyr Glu Gly Asn Lys Trp Asp Thr Ser Tyr Cys Ser Asp
65 70 75 80

Gly Pro Ser Cys Ala Ser Lys Cys Cys Ile Asp Gly Ala Asp Tyr Ser
85 90 95

Ser Thr Tyr Gly Ile Thr Thr Ser Gly Asn Ser Leu Asn Leu Lys Phe
100 105 110

Val Thr Lys Gly Gln Tyr Ser Thr Asn Ile Gly Ser Arg Thr Tyr Leu
Page 12

125

Page 13

3123-4008PC.ST25.txt

Gly Gly Met Val Gln Met Gly Lys Ala Leu Ala Gly Pro Met Val Leu
 370 375 380
 Val Met Ser Ile Trp Asp Asp His Ala Val Asn Met Leu Trp Leu Asp
 385 390 395 400
 Ser Thr Trp Pro Ile Asp Gly Ala Gly Lys Pro Gly Ala Glu Arg Gly
 405 410 415
 Ala Cys Pro Thr Thr Ser Gly Val Pro Ala Glu Val Glu Ala Glu Ala
 420 425 430
 Pro Asn Ser Asn Val Ile Phe Ser Asn Ile Arg Phe Gly Pro Ile Gly
 435 440 445
 Ser Thr Val Ser Gly Leu Pro Asp Gly Gly Ser Gly Asn Pro Asn Pro
 450 455 460
 Pro Val Ser Ser Ser Thr Pro Val Pro Ser Ser Ser Thr Thr Ser Ser
 465 470 475 480
 Gly Ser Ser Gly Pro Thr Gly Gly Thr Gly Val Ala Lys His Tyr Glu
 485 490 495
 Gln Cys Gly Gly Ile Gly Phe Thr Gly Pro Thr Gln Cys Glu Ser Pro
 500 505 510
 Tyr Thr Cys Thr Lys Leu Asn Asp Trp Tyr Ser Gln Cys Leu
 515 520 525

<210> 7
 <211> 4376
 <212> DNA
 <213> Chrysosporium lucknowense

<400> 7
 ggatccacac ctaccatacc ggatagtatg ctacccaagt gacatagggt tggtaaagta 60
 atacgagaac tcagagagca ctgcccataat ggctcgccaa tgacctcaag tgccagggtca 120
 gctttgcgag acagacctga gcgcgtcgga tgtgtgacat ggaacgcgcc ggatcgctt 180
 gttgattaat tatagggaag tagcgaggaa ggtttcagca attgacgtga gcgtacatta 240
 aaagctgtat gatttcagga agacgagcca tggaccagggt ttcaaggctg aatggcttga 300
 cgacttaagc accgaacgag gaatgaaaga atgaaaagtg ggggatcatt ctggcccctc 360
 ctcgtatgtc gagtggttaa gaaggcggtt ctacggagga cctaaagagc tccaatttgc 420
 tctgttgagc ttaagccaca tatctcaaga tgaatacatg tcaggcatag tcaccctgat 480
 cttgttcac agtccacaca cttttcagtt cagcatgttg attcctcatc catatcactt 540

3123-4008PC.ST25.txt

tccattacta tctctttatg tccttgggtca agactccaag gaaccgatag gtgagcatcg	600
gtgaggctcc ctcaaggtag caaagtagcc atcatcaccg aggtctggga atggcgccgt	660
gcccgatctg agtccctcaa ctccacggta cgacgacagc acgtcacatt gacgcaccac	720
ggttgaacaa gcagagaggg acacgtcttg ctacgcgaat cctggcactg gatggagacg	780
cgtgtgagca ggtttccgga accatgacgg cctggtccgg cttctcgaac aaagaagtgg	840
aacacaaaaa gaaccgaaac ggaaacgcag gcacggcatc gacgaccgga ttgtcccacg	900
gggacctcgg ccagtcaagc gttgccctgg ccgtcagctc cctggcgacg gggattcagc	960
acatctcacg ttataggcga cctcatcccc cttccgtctt gtgcggtcgt tgctccgtgc	1020
cgagtaccca ggcgtgccgg ggccttttagc cggggcgga tcagagtcaa gatgcggccg	1080
aattggacgg cagacgaagt ttcgtagagg gtcattgatc gcactgacga caccaccccc	1140
tgcgtgatcc cgtggccctg ggctgggaat tgccggctaa taatctacgg cttaatagat	1200
atgcactttg cacgcggtgc agataaataa gctgtggttt caaacactgg cctccgtact	1260
ttaccacca actgccgctt agcgcgggga cctgagtctt gggagtgcgc ggagcggcag	1320
ccacctcggg ttagcgtaca cacgacggct gcatgcgggg atgccgcgtg catggcttca	1380
tagtgtacga cagaccgtca agtccaaatc tgggtgatgc ttgatgagat gacagcgagc	1440
cccgtcggcg gcaccccggc tatgcatcgc gaattgacaa cactctcagc tctattgcga	1500
cccatcggat aaaagaagaa gaaaaaatg gaccttgagt acgggcgtca gaaacaaaaa	1560
aaaaactccg gaaccaaata tgtcgggcat ggccgggggtg aacgaccgct actccccgtt	1620
cccttcttcg caaacagaac gctacagagg gttttctggt ttgtcaaaga gttcggaggt	1680
cctctgctcc gcgaatgcgt ggtgaaccca ccagcagcca ttgttcttgc atgcgtggcg	1740
gaccgttagc cgctgatcga catggcgagc ttcccacctc agacctggag cagacggttg	1800
cgaggagcaa ggggctgccc tccccctgac ggtcggaccc caatgacttc cccaaacggg	1860
gacatcgagg gtcgtgcatg atggtggaaa gtagttgcag tatgggaagt accccggggt	1920
gccaggaacc gttgttcggc cccccacatt ttctctctgc catgtcaact gtgtgtcgtt	1980
cgagagtcc tggctccggc cccccgtcca attccctaac gggaccgcgg ggcacgcct	2040
gtaactaact tccaaatgaa gccggatatg agggagggag attggatctg gcaagccagc	2100
cattcgctgc gatcggcact cgtccgtcag ccccgagtc catatcccca aaggcaactg	2160
ctcggcgcggt ctcaagtctt cttcggaacg tccagcccga aggcgcgcgc cagcaccggc	2220
cctatgttcc tgattgcgat cctcgatctc cagagacggg tcacctcgcc tcgaggacgg	2280
tgcaggggca tcggcttcgc ttcctagagc tccgggctgt gtgtggtcaa ggggagaagg	2340
cggcggcgcc aagggtgcgtc tcgggcgact caccatcgc ctttaccccc ctcccccca	2400

3123-4008PC.ST25.txt

gtatataaaa gatggccatc gtctcctcgt ctgcttggga agaaaggatc tctcgaccat	2460
gcaccacagc ctagctctaa cccagcttgt cgtgtgttgt tgcccagcat gaagttcgtg	2520
cagtccgcca ccctggcggt cgccgccacg gccctcgctg cgccctcgcg cacgactccc	2580
cagaagcccc gccaggcctc ggcgggctgc gcgtcggccg tgacgctcga tgccagcacc	2640
aacgtgttcc agcagtacac gctgcacccc aacaacttct accgtgccga ggctcgaggct	2700
gccgccgagg ccatctccga ctcggcgctg gccgagaagg cccgcaagggt cgccgacgtc	2760
ggtaccttcc tgtggctcga caccatcgag aacattggcc ggctggagcc cgcgctcgag	2820
gacgtgccct gcgagaacat cgtgggtctc gtcacttacg acctcccggg ccgtgactgc	2880
gcggccaagg cctccaacgg cgagctcaag gtcggcgagc tcgacaggta caagaccgag	2940
tacatcgaca gtgagttaac cttttgtggc cctttctttt ccccgagag agcgtctggt	3000
tgagtgggggt tgtgagagag aaaatggggc gagcttaag actgacgtgt tggctcgag	3060
agatcgccga gatcctcaag gccactcca acacggcctt cgccctcgtc atcgagcccg	3120
actcgtccc caacctggtc accaatagcg acctgcagac gtgccagcag agcgcttccg	3180
gctaccgca ggggtgtcgc tatgccctca agcagctcaa cctccccaac gtggtcatgt	3240
acatcgatgc cggccacggt ggctggctcg gctgggacgc caacctcaag cccggcgccc	3300
aggagctcgc cagcgtctac aagtctgctg gttcgccctc gcaagtccgc ggtatctcca	3360
ccaacgtggc tggttggaac gcctggtaag acactctatg tccccctcgt cggtaaatgg	3420
cgagcggaat ggcgtgaaat gcatggtgct gacctttgat cttttcccc tcctataggg	3480
accaggagcc cggtaggttc tcggacgcct cggatgccca gtacaacaag tgccagaacg	3540
agaagatcta catcaacacc tttggcgctg agtcaagtc tgccggcatg cccaaccacg	3600
ccatcatcga cactggccgc aacggtgtca ccggtctccg cgacgagtgg ggtgactggt	3660
gcaacgtcaa cggcgccggc ttcggtgtgc gcccgactgc caacactggc gacgagctcg	3720
ccgacgcctt cgtgtgggtc aagcccgggt gcgagtccga cggcaccagc gactcgtcgg	3780
cggcgcgcta cgacagcttc tgcggaagc ccgacgcctt caagcccagc cccgaggccg	3840
gtacctggaa ccaggcctac ttcgagatgc tcctcaagaa cgccaaccg tccttctaag	3900
ctctcgacg gcttcttgct gtcagtcgt ctgacgggtg tgtgctggtg gtgcccctgc	3960
tcctgctgct gctgctccgc ggggagggga ggcaacgaaa atgaagtcct gcttcaaac	4020
aaaacagaaa caagcgaggc gcggtgcaat ggtcgtgctg tcgtcttttt tcatgttccc	4080
ttctagtgtg tagttttgat agtcgtacat aaggggtttc agaaccgtct ctctgtctcg	4140
gtctttttgc gagttgttgc gactcgtgat tatggccttt gttgctcgtt gcggcagagt	4200
agaaccacag cgtgttgggg tagcagcttg ctccgtagga cgtagggaaa caacctgaga	4260
ctctggaatt gcagtcagcc tgcgtcgccc ctctaggaaa cgaaggggag aaccagtagt	4320

3123-4008PC.ST25.txt

ggctgcagct tacaaacgcg agcatggtga acatctccga gaaaagggag ggatcc 4376

<210> 8
 <211> 395
 <212> PRT
 <213> Chrysosporium lucknowense

<400> 8

Met Lys Phe Val Gln Ser Ala Thr Leu Ala Phe Ala Ala Thr Ala Leu
 1 5 10 15

Ala Ala Pro Ser Arg Thr Thr Pro Gln Lys Pro Arg Gln Ala Ser Ala
 20 25 30

Gly Cys Ala Ser Ala Val Thr Leu Asp Ala Ser Thr Asn Val Phe Gln
 35 40 45

Gln Tyr Thr Leu His Pro Asn Asn Phe Tyr Arg Ala Glu Val Glu Ala
 50 55 60

Ala Ala Glu Ala Ile Ser Asp Ser Ala Leu Ala Glu Lys Ala Arg Lys
 65 70 75 80

Val Ala Asp Val Gly Thr Phe Leu Trp Leu Asp Thr Ile Glu Asn Ile
 85 90 95

Gly Arg Leu Glu Pro Ala Leu Glu Asp Val Pro Cys Glu Asn Ile Val
 100 105 110

Gly Leu Val Ile Tyr Asp Leu Pro Gly Arg Asp Cys Ala Ala Lys Ala
 115 120 125

Ser Asn Gly Glu Leu Lys Val Gly Glu Leu Asp Arg Tyr Lys Thr Glu
 130 135 140

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

Phe Ala Leu Val Ile Glu Pro Asp Ser Leu Pro Asn Leu Val Thr Asn
 165 170 175

Ser Asp Leu Gln Thr Cys Gln Gln Ser Ala Ser Gly Tyr Arg Glu Gly
 180 185 190

Val Ala Tyr Ala Leu Lys Gln Leu Asn Leu Pro Asn Val Val Met Tyr
 195 200 205

3123-4008PC.ST25.txt

Ile Asp Ala Gly His Gly Gly Trp Leu Gly Trp Asp Ala Asn Leu Lys
 210 215 220

Pro Gly Ala Gln Glu Leu Ala Ser Val Tyr Lys Ser Ala Gly Ser Pro
 225 230 235 240

Ser Gln Val Arg Gly Ile Ser Thr Asn Val Ala Gly Trp Asn Ala Trp
 245 250 255

Asp Gln Glu Pro Gly Glu Phe Ser Asp Ala Ser Asp Ala Gln Tyr Asn
 260 265 270

Lys Cys Gln Asn Glu Lys Ile Tyr Ile Asn Thr Phe Gly Ala Glu Leu
 275 280 285

Lys Ser Ala Gly Met Pro Asn His Ala Ile Ile Asp Thr Gly Arg Asn
 290 295 300

Gly Val Thr Gly Leu Arg Asp Glu Trp Gly Asp Trp Cys Asn Val Asn
 305 310 315 320

Gly Ala Gly Phe Gly Val Arg Pro Thr Ala Asn Thr Gly Asp Glu Leu
 325 330 335

Ala Asp Ala Phe Val Trp Val Lys Pro Gly Gly Glu Ser Asp Gly Thr
 340 345 350

Ser Asp Ser Ser Ala Ala Arg Tyr Asp Ser Phe Cys Gly Lys Pro Asp
 355 360 365

Ala Phe Lys Pro Ser Pro Glu Ala Gly Thr Trp Asn Gln Ala Tyr Phe
 370 375 380

Glu Met Leu Leu Lys Asn Ala Asn Pro Ser Phe
 385 390 395

<210> 9

<211> 5777

<212> DNA

<213> Chrysosporium lucknowense

<400> 9

tgctgctctg atgtgctgat gcacagcttc ccctcgcgat tgccggcagg atctccaacc 60

ctctggatcg gagcagacga tcagcgggca caatggccag cttgccagcg ttcaactcca 120

agttgaccgc cttttatcac gcccaagctg gacatgcaca ggcttggctt ctcgtgttcc 180

tacgatctgc acagtaggtt tgactgctga tcttcgcttt cctgtgcgcc ctccccctcc 240

ctcacgggta ccttatcctt gcctgtaacc ccgcgttatg tcaaacttga gtttgaccaa 300

3123-4008PC.ST25.txt

tgctagcgca aaagtaccta catagtacta tgtaataagg taggtacata catcagtagg	360
cgtttatcta gtaaatttttg gcttttttgaa actcaattgc tcctctcctc gcctccacct	420
ctgcttggca attgacaacc ctggctgtgc ctagaggtag catcgacgat caatcaaadc	480
taaagtattc gagattgacc tttctgctct aattatatta attatccgca caatgctgta	540
gtcattgact ctcttttcaa gttgccttct cgtttatgta tgtacaatgg gcggtcatgc	600
ttcatgccaa cagatgggtc tatcggaaca atgtttgact ttctggtcgc cccgtcgaac	660
tgttttgatt tcgcacggga agtggtctta ccaaagctaa gtcgactcgt ggagcttcgt	720
aacggccagt gatcgttgat cgcttttggga ggagttgcga tggagcgaga ccggctacga	780
gcacgttcgc aaaggcagca cgatagacga ccctccgtgg cgccattcgg gagatgcaca	840
tgacataagc atatcaatac tcacctgaac tcatcggccg atgcctcgca ggtagttaca	900
agacatatct gtgtgggtat attatcccaa cccgtacctt tgtcgcgtca tttcggtatg	960
tgctgatgcc tacttaggga gcaaagacgc ctctcctcac ctgcgggtta cttacttact	1020
gtgcagcatg gccttatgtt ctcccgggtc ttgcttgcgc gaatgaacaa aaacgcccga	1080
agaaaagccg cttcttcgag ttgtgtctac ccgaacataa gaggttattg tcgcagaccg	1140
ccagcaaatg tcaacaaccc acccacggcg ttccagaacc ttcgaaatat catctagttt	1200
aagtttaaat gacggcccga gtcccagccg agattcccat attggccgat accagcggtc	1260
ccttggtttt ccaaggttgt ctctgcaact ggcgcatctg cctacaacga gatataatta	1320
ccgttttctt ttgcaaaagg gcatgcatgg atgtatatta tttatgcctg cagaacgaga	1380
agcaatcatg gtgtaggttt tgtgcggtat ggagctaata atattgaacg gatctctggt	1440
ccgtcctaaa tcgttgaaac gctaggccca ggaggacctg ctcgacttgg cgaacggaga	1500
tttccaggat gaaaggctcg aacatgtcca tccgcgcca gcctgaacac ttttgctcgt	1560
ttccggacca tcgaccacg aaaacagtgc ggttgctggc acagtcagca ctcacgatgg	1620
cgatggtcca gcccgttccc gcccgatgcc cacttgcagc gcaactctcc ttcattcggc	1680
ggccccggcg tgtctggcct attagtacga ttttgatac cggcttggtc gccgcccgcg	1740
tttttcttgg ccgatacggg aatctcggtg gtcccaactc cacctgggca cgctctggtg	1800
ccaacatgga acttcgggat gccgctccg gcacagtcaa gcgctttaa atacgacttt	1860
accccacaag aatcgaggcg taaccggaa ttagggacac ctggacggcg caaccctgg	1920
accgaagggc ctcgctaacc gggttcctgg agccgatgc gcggctgccc gcttgcccgc	1980
tcttgagatg acacttctt tcagcgaggg atggtcgggc agggaaatga tgtattataa	2040
gaagcgagcc gattccgaag gactcgaccc cctctctcgc cctgtgtccg ccagctaatt	2100
acagcactcc ttctcgactt gaaacgccc agatgaagtc ctccatcctc gccagcgtct	2160

3123-4008PC.ST25.txt

tcgccacggg cgccgtggct caaagtgggc cgtggcagca atgtggtggc atcggatggc	2220
aaggatcgac cgactgtgtg tcgggttacc actgctcta ccagaatgat tggtagagcc	2280
agtgcgtgcc tggcgcggcg tcgacaacgc tccagacatc taccacgtcc aggcccaccg	2340
ccaccagcac cgcccctccg tcgtccacca cctcgcttag caagggcaag ctcaagtggc	2400
tcggcagcaa cgagtcgggc gccgagttcg gggagggcaa ctaccccggc ctctggggaa	2460
agcacttcat cttcccgctc acttcggcga ttcaggtacg ggccaataat aatatattat	2520
tatagcaggc aggagggagc aggagaagaa gggaggggca ggtggccaac aatcggaaga	2580
agaccgggag gcaactgacc ttgattcctt tgtgtaatag acgctcatca atgatggata	2640
caacatcttc cggatcgact tctcgatgga gcgtctggtg cccaaccagt tgacgtcgtc	2700
cttcgacgag ggctacctcc gcaacctgac cgaggtggc aacttcgtga cgaacgcggg	2760
caagtacgcc gtcttgacc cgcacaacta cggccggtac tacggcaacg tcatcacgga	2820
cacgaacgcg ttccggacct tctggaccaa cctggccaag cagttcgct ccaactcgct	2880
cgatcatcttc gacaccaaca acgagtacaa cacgatggac cagaccctgg tgctcaacct	2940
caaccaggcc gccatcgacg gcatccgggc cgccggcgcg acctcgagc acatcttcgt	3000
cgagggcaac gcgtggagcg gggcctggag ctggaacacg accaacacca acatggccgc	3060
cctgacggac ccgcagaaca agatcgtgta cgagatgcac cagtagctcg actcggacag	3120
ctcgggcacc cagccgagc gcgtcagcag caacatcggc gccagcgcg tcgtcggagc	3180
caccagtggt ctccgcgcca acggcaagct cggcgctctc ggcgagttcg ccggcgcgcg	3240
caacgccgtc tgccagcagg ccgtcaccgg cctcctcgac cacctccagg acaacagcga	3300
ggtctggctg ggtgccctct ggtgggcccgc cggctcctgg tggggcgact acatgtactc	3360
gttcggtaag tttctccctt gttcttggct tttccccag taaggagtc aggcaacatg	3420
cccaagaccg gctcggcttc gttcaaggc gttcgttgta cactgaag agttccaact	3480
tccaacctg ttcgtgtcct ccgatcagct tcgacggggt gaagggggaa gggatttggg	3540
agtgaggtgg aggtcaaaag gagggatctc ccagatctc cacaacggc cctgagccaa	3600
caacagcctc tgggggtcaaa atgggcgcca accatacggc cattcactca ggacacctgc	3660
taacgcgtct cttttttttg tttccagagc ctcttcggg caccggctat gtcaactaca	3720
actcgatcct aaagaagtac ttgccgtaag gggcatgcag caaggtcgag cgagcattat	3780
tcagggccat ctgcttgtgt cggcaggcat cacgtcaacc catcgaatcg gacagcggaa	3840
tgctccgaga tgccatacac taagtctggt gatgacgtga gaatgctggc cctggtcggg	3900
ggttaccgcc aacaaaaagc acccgagcgc tgccgcgccc ggataccatg gtttcatgta	3960
catattggtt ctttgctttc ttacggggggg gggggggggg gggggctctg cagcgttgct	4020
gagcgattcg tttccaagta tatactttgt ctggaattga attttgagt acattgaccc	4080

3123-4008PC.ST25.txt

aatcaaccag ctcggtgtgc tcacctcccg ttaccccccc tcttctcccc ctgctcggct 4140
tggctttcct ctccggtgtg gagcacggcc acggcgggtcc caatccatat aagatcgatg 4200
gtatactatg gtatacacta gcttggaat aaactaatcc atacgctaac taatggacgg 4260
attatcctaa gggtcaccgg ctcaccgttg gatataacac ctaggatacg ggagagctga 4320
tagaaagggg tgtactccgt attgtactgt acaatacaaa gtacagatag cacacgaagt 4380
acggtaggtg gtcccgcccta gtccggacca acaatagaac atgcgttcct ggggacctgc 4440
aggaaagaag gggggggggg ttgccaagac gcccggggtt caaagaaagc cccggggccgc 4500
cgatgagatg agacggacgc cggcccaagg agaggccggt ggtcgatcct gcaaatgcca 4560
gcaaaaaaaaa tccataccat aatccagtca actttcgtca cactcctgtg aaacgagctg 4620
gagggactgc tggaaggtt ttgcaggta atcactgtat gtggagcatg ccgtacctac 4680
tgtgcttcgt taacagatag agttccagtt gaacacacaa agttctgccc cgcctgccag 4740
acgtgaaaag aagctcctcc gggggagctt taggcaactg ggagggctct ctcccaggtt 4800
catggtgtct gctcttcttc aaatttttat gctgccaccc catttgacag aggtgtgcac 4860
accgttgcca ggtcttgcca tccggcaaaa agcagaaaag tcgacccatc gcctaagaaa 4920
ggcggtcgga aggggatcgg atgctcattg cggcttagcg tctgccatt ctgacgctgc 4980
ccattgtttt gtgtcgcatt cgtcttcgga tgtcggatca agagtcccg attttttccc 5040
ctgtgcttcc agcctaattc gagcgggagc tggctcgggt tcgagtggag ttgccttggt 5100
ggtggagcag caaccagcca attcactccc ccgcattttc gcggccgccc aggcattccc 5160
ggcatgcgtt tgggcggtaa ctactccgta ctggggtagg tgaaattggt tctcccgctg 5220
caggaggctc gtgctcggtc aggggagaac aaagtccaac tgctccttcc tggcaacaat 5280
gagagggggg tctattgcca acgttgacg aaaggagcag ccacaaaacc caaaagcagg 5340
ttaccttact gtacctgagc ttgaacgtcg cgtagcattg gagctctcgt ctaccggcgg 5400
cgtcacactc cattggcagg tcaaggcagt cagtggcagc gaccaacaa cgtcaatgct 5460
tgttacccca gaattacccc gggctgcaac actgcagggg ccgccgccga tgttgatcac 5520
cggttgatta cttctcggcc cgcaaccggg agatgagaag cagaactttg ttctcctttc 5580
aaaaaggacc tgacttgcg ggaacgcact gccggcagtg gagtggatgc acgctagtta 5640
tatgtttccc gccatcccca gtccgcccgt cgcgtccgtg aggctcagtt tggcttcccg 5700
tgccgccgac aaacgagcgg tgcataatta ctttcgctc catgtaccgt gcaccctccc 5760
cgttcgcgac cgtagta 5777

<210> 10
<211> 389
<212> PRT

3123-4008PC.ST25.txt

<213> Chrysosporium lucknowense

<400> 10

```

Met Lys Ser Ser Ile Leu Ala Ser Val Phe Ala Thr Gly Ala Val Ala
1      5      10      15
Gln Ser Gly Pro Trp Gln Gln Cys Gly Gly Ile Gly Trp Gln Gly Ser
20      25
Thr Asp Cys Val Ser Gly Tyr His Cys Val Tyr Gln Asn Asp Trp Tyr
35      40      45
Ser Gln Cys Val Pro Gly Ala Ala Ser Thr Thr Leu Gln Thr Ser Thr
50      55      60
Thr Ser Arg Pro Thr Ala Thr Ser Thr Ala Pro Pro Ser Ser Thr Thr
65      70      75      80
Ser Pro Ser Lys Gly Lys Leu Lys Trp Leu Gly Ser Asn Glu Ser Gly
85      90      95
Ala Glu Phe Gly Glu Gly Asn Tyr Pro Gly Leu Trp Gly Lys His Phe
100     105     110
Ile Phe Pro Ser Thr Ser Ala Ile Gln Thr Leu Ile Asn Asp Gly Tyr
115     120     125
Asn Ile Phe Arg Ile Asp Phe Ser Met Glu Arg Leu Val Pro Asn Gln
130     135     140
Leu Thr Ser Ser Phe Asp Glu Gly Tyr Leu Arg Asn Leu Thr Glu Val
145     150     155     160
Val Asn Phe Val Thr Asn Ala Gly Lys Tyr Ala Val Leu Asp Pro His
165     170     175
Asn Tyr Gly Arg Tyr Tyr Gly Asn Val Ile Thr Asp Thr Asn Ala Phe
180     185     190
Arg Thr Phe Trp Thr Asn Leu Ala Lys Gln Phe Ala Ser Asn Ser Leu
195     200     205
Val Ile Phe Asp Thr Asn Asn Glu Tyr Asn Thr Met Asp Gln Thr Leu
210     215     220
Val Leu Asn Leu Asn Gln Ala Ala Ile Asp Gly Ile Arg Ala Ala Gly
225     230     235     240

```

3123-4008PC.ST25.txt

Ala Thr Ser Gln Tyr Ile Phe Val Glu Gly Asn Ala Trp Ser Gly Ala
 245 250 255

Trp Ser Trp Asn Thr Thr Asn Thr Asn Met Ala Ala Leu Thr Asp Pro
 260 265 270

Gln Asn Lys Ile Val Tyr Glu Met His Gln Tyr Leu Asp Ser Asp Ser
 275 280 285

Ser Gly Thr His Ala Glu Cys Val Ser Ser Asn Ile Gly Ala Gln Arg
 290 295 300

Val Val Gly Ala Thr Gln Trp Leu Arg Ala Asn Gly Lys Leu Gly Val
 305 310 315 320

Leu Gly Glu Phe Ala Gly Gly Ala Asn Ala Val Cys Gln Gln Ala Val
 325 330 335

Thr Gly Leu Leu Asp His Leu Gln Asp Asn Ser Glu Val Trp Leu Gly
 340 345 350

Ala Leu Trp Trp Ala Ala Gly Pro Trp Trp Gly Asp Tyr Met Tyr Ser
 355 360 365

Phe Glu Pro Pro Ser Gly Thr Gly Tyr Val Asn Tyr Asn Ser Ile Leu
 370 375 380

Lys Lys Tyr Leu Pro
 385

<210> 11
 <211> 6060
 <212> DNA
 <213> Chrysosporium lucknowense

<400> 11
 ccggcctcca gttccaggag cttggctctg ccgacatact gtgtacacta ggaattctct 60
 tatgcgggggt gtgcgcgggg aaatgttggg gaactcgagt tgggtcatgt ggacaagacc 120
 aatgggagct gacatcattg tgcgaccctg taaaccggaa gctacaacaa cattctggat 180
 tctacactag tggaagaggt aagtaattga cgacaagcaa gaagcattgc catgttctgc 240
 gaaggatgcg ggtgtttttg catgagcagg aagctgtggc ttttttagtgc tcctttgtgc 300
 tcgccgggcg cgcagaacac taccgaaacg caggggactg cgtgcctctg gggtcgaatg 360
 ccgatcccca tcttcacatt cccaccatcg tgttctgtta acgaagccgg agcggcgggg 420
 actcgaagct ccactacgta tggatacttg ggaccgtacg gagtgtgttg gtacggatgc 480

3123-4008PC.ST25.txt

ctgcacaagt gttgtgcttc ctacgaagac gccaacccac ataatacaca aaagctgttg	540
taagtcgagt tacctcaggc acgttcgggc aactcgggca acctgacgag atttccccgc	600
cattccgcca agaggccggc gcctgccctg attaggcagc tcttggaaca atactatgta	660
gaatggaagc tccatccata gtcagctcca ttggcgggtcc cagtgatctc gatggctgga	720
tggctgctct gtacggtaca tacatagtaa gttctcgcct tgagagccca attcgctgca	780
atagcatctt tccccgcagt gcgcgggccc ccctgggtcc cgctccacaa tgaccttgct	840
tctggagctt ctcgacgaac agatcggccc gtttcttctc cacaccaatc cgaaccagtc	900
gggagcatgg ctgcggatgc gacgcagcct tccttcgcgc tgtacaaaca gctccgggaa	960
cgctgactgg tatgtacgga ctacagtaag tacactacga gtgcacatac tgacgaatac	1020
cggcctcaga ggaacctggc aggaccctac cccacacgaa accacagcga gaaagcgcaa	1080
tggatcagta actactgcga agtaaccgtg gtcccgggca aaggatctga gggccgatcg	1140
ctcgtggggc tgcgaggcga gggagagcaa acaagccagt cctcccgcga acctggaaaa	1200
tcacttataa acacacgtca ccggcgccgg ggtgcgcgcc atgtgtcacc tccaggctcc	1260
tcccgggcca tgatctctgc cgggtgccatc aatcatctcg gttcgccgca gctgcttctt	1320
tctgtgcagt gaacgctctc aaactgcaac gacgctgtcc gacatgaagg ctgctgcgct	1380
ttcctgcctc ttcggcagta cccttgccgt tgcaggcgcc attgaatcga gaaaggatatg	1440
gacgggcttt cgtcaaagac tcgctccccg atcaacttcc cctttcatcc agaccacccc	1500
aaccttccca gtcttgcttc gagcacgata tcttcgggca gcacccacc cacatccact	1560
cagattagcg gcgacaccgt tgactgttgc aatccgcaat cgacatgcaa cttccagccg	1620
cagcccaatg gctgctcacg cttcccgcga aagcctcact tgctgacaat catcgtcagg	1680
ttcaccagaa gcccctcgcg agatctgaac ctttttacc gtcgccatgg atgaatccca	1740
acgccgacgg ctgggaggag gcctatgccc aggccaagtc ctttgtctcc caaatgactc	1800
tgctagagaa ggtcaacttg accacgggag tcgggtaagt tttgtcattt tgtccaggta	1860
acatgcaaat ggttctgcta acaataactt accgtagctg gggggctgag cagtgcgtcg	1920
gccaagtggg cgcgatccct cgccttgac ttcgcagtct gtgcatgcat gactccccctc	1980
tcggcatccg aggagccgac tacaactcag cgttccccctc tggccagacc gttgctgcta	2040
cctgggatcg cggctctgatg tacgctcgcg gctacgcaat gggccaggag gccaaaggca	2100
agggcatcaa tgtccttctc ggaccagtcg ccggccccctc tggccgcatg cccgagggcg	2160
gtcgtaactg ggaaggcttc gctccggatc ccgtccttac cggcatcggc atgtccgaga	2220
cgatcaaggg cattcaggat gctggcgta tcgcttgtgc gaagcacttt attggaaacg	2280
agcaggggtga gtagtcaaag acgggcccgtc tcggacccgc ggcttcaagc tgctgactct	2340
gctgcagagc acttcagaca ggtgccagaa gccaggggat acggttataa catcagcgaa	2400

3123-4008PC.ST25.txt

accctctcct ccaacattga cgacaagacc atgcacgagc tctacctttg gccgttttgcc 2460
 gatgccgtcc gggccggcgt cggctctgtc atgtgctcgt accagcaggt caacaactcg 2520
 tacgcctgcc agaactcgaa gctgctgaac gacctcctca agaacgagct tgggtttcag 2580
 ggcttcgtca tgagcgactg gcaggcacag cacttggcg cagcaagcgc cgtggctggg 2640
 ctcgatatgt ccatgccggg cgacaccag ttcaacactg gcgtcagttt ctggggcgcc 2700
 aatctcacc ctcgccgtcct caacggcaca gtccctgcct accgtctcga cgacatggcc 2760
 atgcgcatca tggccgcctt cttcaaggct accaagacca cccacctgga acccatcaac 2820
 ttctccttct ggaccgacga cacttatggc ccgatccact gggccgcca gcatggctac 2880
 cagaagatta attcccacgt tgacgtccgc gccgaccag gcaacctcat ccgggagatt 2940
 gccgccaagg gtacgggtgt gctgaagaat accggctctc taccctgaa caagccaaag 3000
 ttcgtggccg tcatcggcga ggatgctggg tcgagccca acgggcccga cggctgcagc 3060
 gaccgaggct gtaacgaagg cacgctcgcc atgggctggg gatccggcac agccaactat 3120
 ccgtacctcg tttccccga cgccgcgtc caggcccggg ccatccagga cggcacgagg 3180
 tacgagagcg tcctgtccaa ctacgccgag gaaaagacaa aggtcttggg ctgcaggcc 3240
 aatgcaaccg ccatcgtctt cgtcaatgcc gactcaggcg agggctacat caacgtggac 3300
 ggtaacgagg gcgaccgtaa gaacctgact ctctggaaca acggtgatac tctgggtcaag 3360
 aacgtctcga gctggtgcag caacaccatc gtcgtcatcc actcggtcgg cccggtcctc 3420
 ctgaccgatt ggtacgacaa cccaacatc acggccattc tctgggctgg tcttccgggc 3480
 caggagtcgg gcaactccat caccgacgtg ctttacggca aggtcaacct cgccgccgc 3540
 tcgcccttca cttggggcaa gaccgcgaa agctatggcg cggacgtcct gtacaagccg 3600
 aataatggca atggtgcgcc ccaacaggac ttcaccgagg gcgtcttcat cgactaccgc 3660
 tacttcgaca aggttgacga tgactcgggtc atctacgagt tcggccacgg cctgagctac 3720
 accaccttcg agtacagcaa catccgcgtc gtcaagtcca acgtcagcga gtaccggccc 3780
 acgacgggca ccacggccca ggccccgacg tttggcaact tctccaccga cctcaggagc 3840
 tatctcttcc ccaaggacga gttcccctac atctaccagt acatctacct gtacctcaac 3900
 acgaccgacc cccggagggc ctcggccgat cccactacg gccagaccgc cgaggagtgc 3960
 ctcccgcccc acgccaccga tgacgacccc cagccgctcc tccggtcctc gggcggaac 4020
 tccccggcg gcaaccgcca gctgtacgac attgtctaca caatcacggc cgacatcacg 4080
 aatacgggct ccgttgtagg cgaggaggta ccgcagctct acgtctcgtt gggcggtccc 4140
 gaggatccca aggtgcagct gcgcgacttt gacaggatgc ggatcgaacc cggcgagacg 4200
 aggcagttca ccggccgcct gacgcgcaga gatctgagca actgggacgt cacgggtgcag 4260

3123-4008PC.ST25.txt

gactgggtca	tcagcaggta	tcccaagacg	gcatatgttg	ggaggagcag	ccggaagttg	4320
gatctcaaga	ttgagcttcc	ttgaatgagt	ttcatcaggg	gctgcagagg	gatggtaaca	4380
cgttcttaat	cagaagtatg	atggagaaaa	gcacttggca	agttccggtg	agcaaaaaga	4440
aggcacttat	taagtgtagg	gcggtgttct	atgtttaata	ggtgctatgt	ttacatataa	4500
ttagtatata	atgatttaaat	aattatgttt	agcagttgct	aatgtcgtaa	atttcggcgt	4560
gtgatgactg	ctacaacact	ggttctgtct	tctagtcgcc	attgttaatt	atgaaggtta	4620
ttgtctacaa	tttctaatac	cttatggatg	attgcccagc	tggtttcaaa	ctcgttacgc	4680
gcaaattggt	cgattgaggt	attattcatt	gtaagtacct	ccgtacagcg	tccccaacta	4740
tttccattca	cgagatgcct	cgcttttcgg	tgctttcggg	acagggctgg	cagcggatca	4800
tggcgcgatc	aaaacatggc	gagcagctgt	ccaggacgga	ggacaggttg	gggactgatg	4860
cctcccgga	gcattaaggt	cagaagatag	acacgtttta	cacagcgttg	agaccgacaa	4920
gccacattag	gcagcgccgg	ttgcaccacc	gccgtcacgg	gcaacggttc	aatcaatcga	4980
caacagtgga	agacaaagta	ctgaagatca	ggtattaata	gtgtgagaga	gaaacagacg	5040
gtggaactag	ggtgctaata	tttctcttga	tttcggtgtc	catggtagta	cagaacacaa	5100
gaaaaagaag	gaggagtga	cggagaagga	ggagggggaa	gccagaaaaa	agaacatgaa	5160
aaagcataca	cattggagtc	ggtcagtcgg	ttgattgggt	tggtagagag	cgaaaaagca	5220
agcgtcacct	gtaggattcg	aacctacgt	cccgaaggaa	ctgcctaaga	acgctaagca	5280
aggttagcag	ggcagcgcg	taaccactcc	gccaaagtga	ctgtcgttga	tcatggtcga	5340
attcaagtag	cttataggag	ttcaaccaga	tcacaaatgc	ataggtgctc	gtagaacggt	5400
ctaagtatga	gttgattata	agcaaccgaa	tggctctcag	cggcaacacc	gtagctgaag	5460
taacaaaacg	cacctttgg	tactttctga	ctataaaaat	gggatatttg	gaaatgacca	5520
cccgataagg	tgtcaaatc	taaatgactg	tctgggtgtg	aagatgttac	tgtggttcca	5580
ccacgaacca	gttttagtat	ccgcatgctt	cagtctctgc	gcctcgacag	gcggagggtg	5640
tgtgttagat	cagaatcgat	gtgacgctgt	gaccgcgagg	ctctcgagcc	taggtgcggt	5700
agttctgttc	aaaaagaagt	gtgtggccgg	gtttgggctg	ccttatagcc	taccatcctg	5760
gctgtgggtc	ccgagcggga	gccggttctc	cgttttgggt	ccgataaagt	gtcatatctg	5820
cctcccggtt	tcgcatctaa	tttctgactt	cgttcgggac	ctctggagac	gtagggatag	5880
gtatgggata	tgcccgcat	ttcgtaaatg	tccatagtct	ctttcgggac	gaggcgga	5940
gctctcagag	ctatctaagc	ttaaccaacc	cctgatcctt	aaccctccca	gaccacacct	6000
cctgggagaa	taaaccgggc	tccaagatcg	aatcgaaat	cagtgcgcga	acttgaaatc	6060

<210> 12
<211> 871

3123-4008PC.ST25.txt

<212> PRT

<213> Chrysosporium lucknowense

<400> 12

Met Gln Leu Pro Ala Ala Ala Gln Trp Leu Leu Thr Leu Pro Ala Lys
 1 5 10 15

Ala Ser Leu Ala Asp Asn His Arg Gln Val His Gln Lys Pro Leu Ala
 20 25 30

Arg Ser Glu Pro Phe Tyr Pro Ser Pro Trp Met Asn Pro Asn Ala Asp
 35 40 45

Gly Trp Ala Glu Ala Tyr Ala Gln Ala Lys Ser Phe Val Ser Gln Met
 50 55 60

Thr Leu Leu Glu Lys Val Asn Leu Thr Thr Gly Val Gly Trp Gly Ala
 65 70 75 80

Glu Gln Cys Val Gly Gln Val Gly Ala Ile Pro Arg Leu Gly Leu Arg
 85 90 95

Ser Leu Cys Met His Asp Ser Pro Leu Gly Ile Arg Gly Ala Asp Tyr
 100 105 110

Asn Ser Ala Phe Pro Ser Gly Gln Thr Val Ala Ala Thr Trp Asp Arg
 115 120 125

Gly Leu Met Tyr Arg Arg Gly Tyr Ala Met Gly Gln Glu Ala Lys Gly
 130 135 140

Lys Gly Ile Asn Val Leu Leu Gly Pro Val Ala Gly Pro Leu Gly Arg
 145 150 155 160

Met Pro Glu Gly Gly Arg Asn Trp Glu Gly Phe Ala Pro Asp Pro Val
 165 170 175

Leu Thr Gly Ile Gly Met Ser Glu Thr Ile Lys Gly Ile Gln Asp Ala
 180 185 190

Gly Val Ile Ala Cys Ala Lys His Phe Ile Gly Asn Glu Gln Glu His
 195 200 205

Phe Arg Gln Val Pro Glu Ala Gln Gly Tyr Gly Tyr Asn Ile Ser Glu
 210 215 220

Thr Leu Ser Ser Asn Ile Asp Asp Lys Thr Met His Glu Leu Tyr Leu
 225 230 235 240

3123-4008PC.ST25.txt

Trp Pro Phe Ala Asp Ala Val Arg Ala Gly Val Gly Ser Val Met Cys
 245 250 255

Ser Tyr Gln Gln Val Asn Asn Ser Tyr Ala Cys Gln Asn Ser Lys Leu
 260 265 270

Leu Asn Asp Leu Leu Lys Asn Glu Leu Gly Phe Gln Gly Phe Val Met
 275 280 285

Ser Asp Trp Gln Ala Gln His Thr Gly Ala Ala Ser Ala Val Ala Gly
 290 295 300

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

Phe Trp Gly Ala Asn Leu Thr Leu Ala Val Leu Asn Gly Thr Val Pro
 325 330 335

Ala Tyr Arg Leu Asp Asp Met Ala Met Arg Ile Met Ala Ala Leu Phe
 340 345 350

Lys Val Thr Lys Thr Thr His Leu Glu Pro Ile Asn Phe Ser Phe Trp
 355 360 365

Thr Asp Asp Thr Tyr Gly Pro Ile His Trp Ala Ala Lys His Gly Tyr
 370 375 380

Gln Lys Ile Asn Ser His Val Asp Val Arg Ala Asp His Gly Asn Leu
 385 390 395 400

Ile Arg Glu Ile Ala Ala Lys Gly Thr Val Leu Leu Lys Asn Thr Gly
 405 410 415

Ser Leu Pro Leu Asn Lys Pro Lys Phe Val Ala Val Ile Gly Glu Asp
 420 425 430

Ala Gly Ser Ser Pro Asn Gly Pro Asn Gly Cys Ser Asp Arg Gly Cys
 435 440 445

Asn Glu Gly Thr Leu Ala Met Gly Trp Gly Ser Gly Thr Ala Asn Tyr
 450 455 460

Pro Tyr Leu Val Ser Pro Asp Ala Ala Leu Gln Ala Arg Ala Ile Gln
 465 470 475 480

Asp Gly Thr Arg Tyr Glu Ser Val Leu Ser Asn Tyr Ala Glu Glu Lys

3123-4008PC.ST25.txt
490

485

495

Thr Lys Ala Leu Val Ser Gln Ala Asn Ala Thr Ala Ile Val Phe Val
500 505 510

Asn Ala Asp Ser Gly Glu Gly Tyr Ile Asn Val Asp Gly Asn Glu Gly
515 520 525

Asp Arg Lys Asn Leu Thr Leu Trp Asn Asn Gly Asp Thr Leu Val Lys
530 535 540

Asn Val Ser Ser Trp Cys Ser Asn Thr Ile Val Val Ile His Ser Val
545 550 555 560

Gly Pro Val Leu Leu Thr Asp Trp Tyr Asp Asn Pro Asn Ile Thr Ala
565 570 575

Ile Leu Trp Ala Gly Leu Pro Gly Gln Glu Ser Gly Asn Ser Ile Thr
580 585 590

Asp Val Leu Tyr Gly Lys Val Asn Pro Ala Ala Arg Ser Pro Phe Thr
595 600 605

Trp Gly Lys Thr Arg Glu Ser Tyr Gly Ala Asp Val Leu Tyr Lys Pro
610 615 620

Asn Asn Gly Asn Gly Ala Pro Gln Gln Asp Phe Thr Glu Gly Val Phe
625 630 635 640

Ile Asp Tyr Arg Tyr Phe Asp Lys Val Asp Asp Asp Ser Val Ile Tyr
645 650 655

Glu Phe Gly His Gly Leu Ser Tyr Thr Thr Phe Glu Tyr Ser Asn Ile
660 665 670

Arg Val Val Lys Ser Asn Val Ser Glu Tyr Arg Pro Thr Thr Gly Thr
675 680 685

Thr Ala Gln Ala Pro Thr Phe Gly Asn Phe Ser Thr Asp Leu Glu Asp
690 695 700

Tyr Leu Phe Pro Lys Asp Glu Phe Pro Tyr Ile Tyr Gln Tyr Ile Tyr
705 710 715 720

Pro Tyr Leu Asn Thr Thr Asp Pro Arg Arg Ala Ser Ala Asp Pro His
725 730 735

3123-4008PC.ST25.txt

Tyr Gly Gln Thr Ala Glu Glu Phe Leu Pro Pro His Ala Thr Asp Asp
 740 745 750

Asp Pro Gln Pro Leu Leu Arg Ser Ser Gly Gly Asn Ser Pro Gly Gly
 755 760 765

Asn Arg Gln Leu Tyr Asp Ile Val Tyr Thr Ile Thr Ala Asp Ile Thr
 770 775 780

Asn Thr Gly Ser Val Val Gly Glu Glu Val Pro Gln Leu Tyr Val Ser
 785 790 795 800

Leu Gly Gly Pro Glu Asp Pro Lys Val Gln Leu Arg Asp Phe Asp Arg
 805 810 815

Met Arg Ile Glu Pro Gly Glu Thr Arg Gln Phe Thr Gly Arg Leu Thr
 820 825 830

Arg Arg Asp Leu Ser Asn Trp Asp Val Thr Val Gln Asp Trp Val Ile
 835 840 845

Ser Arg Tyr Pro Lys Thr Ala Tyr Val Gly Arg Ser Ser Arg Lys Leu
 850 855 860

Asp Leu Lys Ile Glu Leu Pro
 865 870

<210> 13
 <211> 998
 <212> DNA
 <213> Chrysosporium lucknowense

<400> 13
 atggcccatc tctccgccac caccgggttc ctgcacctcc cggccctggc cctggccccag 60
 ctctcgggca gcgccagac gaccgggtac tgggactgct gcaagccgag ctgcgcctgg 120
 cccggcaagg gccctcgtc tccggtgcag gcctgcgaca agaacgacaa cccgctcaac 180
 gacggcggct ccacccggtc cggctgcgac gcgggaggca gcgcctacat gtgctcctcc 240
 cagagccccct gggccgtcag cgacgagctg tcgtacggct gggcggccgt caagctcgcc 300
 ggcagctccg agtcgcagtg gtgctgcgcc tgctacgagc tgaccttcac cagcggggccg 360
 gtcgcgggca agaagatgat tgtgcaggcg accaacaccg gtggcgacct gggcgacaac 420
 cactttgacc tggccgtgag ttgcctcccc ttctccccgg accgctcaga ttagatgaga 480
 ttagactttg ctcgtaaadc ggtccaagat tcccttgact gaccaacaaa catcatacgg 540
 gcagatcccc ggtggcggtg tcggtatttt caacggtaag ctggtgcccc cggaccctc 600
 cccggacccc tccccctttt cctccagcga gccgagttgg gatcgccgag atcgagaact 660

3123-4008PC.ST25.txt

cacacaactt ctctctcgac agcctgcacc gaccagtacg gcgctccccc gaacggctgg 720
 ggcgaccgct acggcgggcat ccattccaag gaagagtgcg aatccttccc ggaggccctc 780
 aagcccggct gcaactggcg cttcgactgg tacgttgctt tgacataccg gaaccaatt 840
 cctccaaccc cccccctttt ctcccccaac tccgggggta gtcggaatgt cgcgactgac 900
 cctatttcag gttccaaaac gccgacaacc cgtcggtcac cttccaggag gtggcctgcc 960
 cgtcggagct cacgtccaag agcggctgct cccgttaa 998

<210> 14
 <211> 225
 <212> PRT
 <213> Chrysosporium lucknowense

<400> 14

Met His Leu Ser Ala Thr Thr Gly Phe Leu Ala Leu Pro Ala Leu Ala
 1 5 10 15

Leu Ala Gln Leu Ser Gly Ser Gly Gln Thr Thr Arg Tyr Trp Asp Cys
 20 25 30

Cys Lys Pro Ser Cys Ala Trp Pro Gly Lys Gly Pro Ser Ser Pro Val
 35 40 45

Gln Ala Cys Asp Lys Asn Asp Asn Pro Leu Asn Asp Gly Gly Ser Thr
 50 55 60

Arg Ser Gly Cys Asp Ala Gly Gly Ser Ala Tyr Met Cys Ser Ser Gln
 65 70 75 80

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

Lys Leu Ala Gly Ser Ser Glu Ser Gln Trp Cys Cys Ala Cys Tyr Glu
 100 105 110

Leu Thr Phe Thr Ser Gly Pro Val Ala Gly Lys Lys Met Ile Val Gln
 115 120 125

Ala Thr Asn Thr Gly Gly Asp Leu Gly Asp Asn His Phe Asp Leu Ala
 130 135 140

Ile Pro Gly Gly Gly Val Gly Ile Phe Asn Ala Cys Thr Asp Gln Tyr
 145 150 155 160

Gly Ala Pro Pro Asn Gly Trp Gly Asp Arg Tyr Gly Gly Ile His Ser
 165 170 175

3123-4008PC.ST25.txt

Lys Glu Glu Cys Glu Ser Phe Pro Glu Ala Leu Lys Pro Gly Cys Asn
 180 185 190

Trp Arg Phe Asp Trp Phe Gln Asn Ala Asp Asn Pro Ser Val Thr Phe
 195 200 205

Gln Glu Val Ala Cys Pro Ser Glu Leu Thr Ser Lys Ser Gly Cys Ser
 210 215 220

Arg
 225

<210> 15
 <211> 4190
 <212> DNA
 <213> Chrysosporium lucknowense

<400> 15
 gcgcttccgg cctgggagag taaaatgacg gaagccgggc cccgtccgac tgcgtttgtc 60
 ccaactcgga agcaggcatc gttttttggg cgggaggaag cgttgcaaca cgcactatcg 120
 ccaaggtgga ctcggcgcaa tctggagggtt cggcccgcgg aggacggaat ccgggctgaa 180
 tctgcgcaaa ggctgaccct gcgatggtgg gaaaatgtaa atatgtgaag ttataggcat 240
 ataggactca gcgatgacat ggaaattgca gaggcatgtg ggatttcagc gtttggcatg 300
 cattggtcgg atctctcgcc ttgtctgatg tgatcccgcc ggaggtgttt cggctctctgg 360
 ggaaggggacc ccccttgcc cccacctgc cccgcacatc gcctcgccac gactcccgcg 420
 cgccgaggaa gaacttcggg tctttgtgac gggagattcc actgagtgag cattggccaa 480
 ccaagcacac aattactccg tacatacaca gtacttctga ctccgtaaag taaaccgtgt 540
 gtttcaaaga tcggtaatcc gtaacaggta ctccgtatct aaggtaaatt taccctgtgc 600
 acggagcaga acctgaactt cttccccctt cttactcgag tagtcaccct actccaacca 660
 gcggcttttc aactcgcaaa gtcttgttta taacagtgc tatacctgca tttcgtatct 720
 cgctagtgtg aagacgacca cacgcggaca aagaaagaaa aatccaattg cccgatggct 780
 cttagtttga ggacagcagc gaaggactac actgcgccgt agtgaccagg ccaagaaacg 840
 cgaatcgtat attaacggca aatcaaaatg gattatatgc catttcgctt ccgggttgcg 900
 tgctcgtccg aagtctggtg ccgatcgatt gcgaaccccc ggaatcgcg gatgattcct 960
 acagccgccg aaaggggggg ggggggaggg ggggtctggac gggacgtgca taacttcgaa 1020
 tttctagaat attgcggatt gggttccctt cagccctgcg agcgcgcccc cttctggaac 1080
 cgcacccttc accggttcca cacacagagg acatgggtgg aaatgtgtac ctgacggttg 1140
 cccctttggg acagtggaga ggcggatgtt cggataacca tccggagccg cagtgtcgac 1200

3123-4008PC.ST25.txt

caagatcttg gcttaccatc gacaccaaca tgcggactcg tccctcagtc atggagccctt 1260
 ggctcgcgga gcctccgttc gaagcggcta tcccgtcctg ccagcggagg atctcgtacc 1320
 gcttccgcga actgtgaatg tcctgggtat aagagcatgg cgcgaccttg tctcgtcagg 1380
 aacggggagg aggagggtt ggtagggtc gcgttcgttt ggagattgct gagctctgag 1440
 ccttcgggtcc ttggatccct gcggtccccg gtctcctctc tctctctctc tctctctctc 1500
 tctctctctt cttcccacgc tcgttcgaca gacgcctccc cttcttcgct ctctttccc 1560
 tcgcacgtag cacactaata gtgcaccatg cgcgtctcta gtttggtcgc ggcccttgct 1620
 accggtggtc ttgtcgccgc cacgcctaag cccaaggggt cgtcgcccc tggggccgtg 1680
 gacgcgaacc ctttcaagg caagacgcag ttcgtcaacc cggcatgggc ggccaagctg 1740
 gaacagacca aaaaggcgtt cctggccagg aacgacaccg tcaatgccgc caagacggag 1800
 aaggtccagc agaccagctc gttcgtctgg gtctcgagga tcgccgagct ctccaacatc 1860
 gacgacgcca tcgcggtgc ccgcaaggcg cagaagaaga cgggcaggag gcagatcgtc 1920
 ggccctggtg tctacaacct tccggaccgc gactgcagcg cgggcgagag cgcgggagag 1980
 ctgagcagcg acaagaacgg gctcgagatc tacaagactg agttcgtcaa gcccttcgcc 2040
 gacaagggtg cgcccgcaaa ggacctcgac ttcgccatcg tcctggagcc cgactcgtg 2100
 gccaacctgg tcaccaacct gggcatcgag ttctgcgcca acgccgccc cgtctaccgc 2160
 gagggcatcg cctatgccat ctccagcctt cagcagccaa acgtgcactt gtacatcgat 2220
 gctgcccacg gcggctggct cggctgggac gacaacctgc cgctggccgc caaggagttt 2280
 gccgaggtg tcaagcttgc cggcgagggc aagaagatcc gcggcttcgt caccaacgtg 2340
 tccaactaca accccttcca cgccgtcgtg cgcgagaact ttaccgagtg gagcaactcg 2400
 tgggacgagt ctactacgc ctctcgtc acaccgttcc tcgagaaaga ggggctgccg 2460
 gcacgcttca tcgtcgacca ggtcgcgtt gccctcccgg gagcccgcaa ggagtgggta 2520
 gtttcgacca gattgaccct cgacccatgc gaccgagatt gctgacgatt gaattgcgtg 2580
 tcccgtcccc caggggtgaa tggtgcaacg tggcaccgc cggttttggc cccgcgcca 2640
 cgaccagggt caacaacacc gtcgtcgatg ctctcgtctg ggtcaagcct ggcggcgaga 2700
 gcgacggcga gtgtggcttg gctggcgccc ccaaggccgg ccagtggttc gacgagtacg 2760
 cccagatgct ggtcgagaat gccacccgt ctgtcgtcca caagtggtag ataaattttg 2820
 gagtccgaga aggggtcccag atagactttt gttttaaaac aaaatgcaag gtgtcgacag 2880
 atactggctt aacattaacc aagcaccatg aacatgactt gtcaacatat tgatacttc 2940
 cgctgctttc ccatacgtgc tctcaggtct cagggatcaa atggataggt cggtaatgca 3000
 aaacgatcca ttggatatcc agaagagaga aaaaaaaaaag gacatgcatg ccttgtctgt 3060

3123-4008PC.ST25.txt

catcatgagg aaacaaagga aaaacaaacg atcgtcgtgt tccaacaagc tttccaagac 3120
cacaagaccc atccaccaac acaaccaaac gacaagcaat acgatggacc gccgttggtc 3180
catctctcaa gagctgacta aacgaacagt cggtgaaatc atcctacatg agtacgccgc 3240
accacctgtt atcgtgtaaa ccaaatcgcc tggttaaagt catcatctct taggtatgat 3300
cgtaagttcc ggtcacggtc acggatcagg gatggttctc aattcgtgtg tcgcgtagcc 3360
gccgccgtat ctggacaaga ctctttgtat tgctccgaaa ccgcttttgc cgccctaata 3420
atctgtagcc ttcttacctg gtggtgcctt gaaagacgcg gcaggcaaca cttcgcaggc 3480
ctgtggcgca ccagcaccag gctgtggtga tgccccggaa ccggtcgtcg acttgctcgc 3540
ggtgtcctcg gctggtgggg atgggggtga tgagggcttg gaggggtgtg ttgcgcccgc 3600
aacatccggc tccggctccg gaccgtccac agacattgga cctgcgagca tgactcgtgc 3660
cttcagccag accaaagcca tgccatcatc gcctctgccg acgctgttga gcgggaggct 3720
gatgttctca gccagaactg cgggctgtac ggccatgacc atgggctgtt cggctctggc 3780
gtcttgccgc ggtttctccc tgccagcttg ttgtgcgcgg tgcctgcgag attcgacttc 3840
gacctgggcg tggcagaggg tgacgagggc cggtgacgcc ttgatctcct tgctcccat 3900
gtccttccac ccgtacaggc ggacgggtgc catacgcgtc cacagcctgc acgagaacct 3960
cagggcgtcg tcaatgagtt ctgtcaactt gctctccagc ctctctatgc cgcgagcatc 4020
ctgatcctgg agcagaaacc gtgccgagcc tccgaggaaa cgctccttca gcttccgcgc 4080
gtagtttagg cgtgattcaa caaacgtccg gcgggactcg ttgttgcccg cagcagcgac 4140
gtccttgatg ctgaagccgc cgtcggcgaa caggcgcac atctgggccc 4190

<210> 16
<211> 381
<212> PRT
<213> Chrysosporium lucknowense

<400> 16

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

Ala Ala Thr Pro Lys Pro Lys Gly Ser Ser Pro Pro Gly Ala Val Asp
20 25 30

Ala Asn Pro Phe Lys Gly Lys Thr Gln Phe Val Asn Pro Ala Trp Ala
35 40 45

Ala Lys Leu Glu Gln Thr Lys Lys Ala Phe Leu Ala Arg Asn Asp Thr
50 55 60

Val Asn Ala Ala Lys Thr Glu Lys Val Gln Gln Thr Ser Ser Phe Val
Page 34

65										70					75					80				
Trp	Val	Ser	Arg	Ile ₈₅	Ala	Glu	Leu	Ser	Asn ₉₀	Ile	Asp	Asp	Ala	Ile ₉₅	Ala									
Ala	Ala	Arg	Lys ₁₀₀	Ala	Gln	Lys	Lys	Thr ₁₀₅	Gly	Arg	Arg	Gln	Ile ₁₁₀	Val	Gly									
Leu	Val	Leu ₁₁₅	Tyr	Asn	Leu	Pro	Asp ₁₂₀	Arg	Asp	Cys	Ser	Ala ₁₂₅	Gly	Glu	Ser									
Ala	Gly ₁₃₀	Glu	Leu	Ser	Ser	Asp ₁₃₅	Lys	Asn	Gly	Leu	Glu ₁₄₀	Ile	Tyr	Lys	Thr									
Glu ₁₄₅	Phe	Val	Lys	Pro	Phe ₁₅₀	Ala	Asp	Lys	Val	Ala ₁₅₅	Ala	Ala	Lys	Asp	Leu ₁₆₀									
Asp	Phe	Ala	Ile	Val ₁₆₅	Leu	Glu	Pro	Asp	Ser ₁₇₀	Leu	Ala	Asn	Leu	Val ₁₇₅	Thr									
Asn	Leu	Gly	Ile ₁₈₀	Glu	Phe	Cys	Ala	Asn ₁₈₅	Ala	Ala	Pro	Val	Tyr ₁₉₀	Arg	Glu									
Gly	Ile	Ala ₁₉₅	Tyr	Ala	Ile	Ser	Ser ₂₀₀	Leu	Gln	Gln	Pro	Asn ₂₀₅	Val	His	Leu									
Tyr	Ile ₂₁₀	Asp	Ala	Ala	His	Gly ₂₁₅	Gly	Trp	Leu	Gly	Trp ₂₂₀	Asp	Asp	Asn	Leu									
Pro ₂₂₅	Leu	Ala	Ala	Lys	Glu ₂₃₀	Phe	Ala	Glu	Val	Val ₂₃₅	Lys	Leu	Ala	Gly	Glu ₂₄₀									
Gly	Lys	Lys	Ile	Arg ₂₄₅	Gly	Phe	Val	Thr	Asn ₂₅₀	Val	Ser	Asn	Tyr	Asn ₂₅₅	Pro									
Phe	His	Ala	Val ₂₆₀	Val	Arg	Glu	Asn	Phe ₂₆₅	Thr	Glu	Trp	Ser	Asn ₂₇₀	Ser	Trp									
Asp	Glu	Ser ₂₇₅	His	Tyr	Ala	Ser	Ser ₂₈₀	Leu	Thr	Pro	Phe	Leu ₂₈₅	Glu	Lys	Glu									
Gly	Leu ₂₉₀	Pro	Ala	Arg	Phe	Ile ₂₉₅	Val	Asp	Gln	Gly	Arg ₃₀₀	Val	Ala	Leu	Pro									
Gly ₃₀₅	Ala	Arg	Lys	Glu	Trp ₃₁₀	Gly	Glu	Trp	Cys	Asn ₃₁₅	Val	Ala	Pro	Ala	Gly ₃₂₀									

3123-4008PC.ST25.txt

Phe Gly Pro Ala Pro Thr Thr Arg Val Asn Asn Thr Val Val Asp Ala
 325 330 335

Leu Val Trp Val Lys Pro Gly Gly Glu Ser Asp Gly Glu Cys Gly Leu
 340 345 350

Ala Gly Ala Pro Lys Ala Gly Gln Trp Phe Asp Glu Tyr Ala Gln Met
 355 360 365

Leu Val Glu Asn Ala His Pro Ser Val Val His Lys Trp
 370 375 380

<210> 17
 <211> 3000
 <212> DNA
 <213> Chrysosporium lucknowense

<400> 17
 cgcgccccg tctttgaacg cttgagaagc gcacggtgaa gaaccatcaa ctccgattcc 60
 gctcctcatc ctcccacgaa gccgattgaa atagccacag cggctatgta cggattactc 120
 tgctccgttt gcacatccat acacagcgct atttttaaaa gttcaggacg gccaagcccc 180
 gttcttgaa cggacgaccc ggattccgaa agctccagcg ctcaatgcgg tcagtcgtgg 240
 cgctgatacct gctgatctgc tgatctcata aaccgcgaac ttcaactttt cactttgaag 300
 cgtatacacg cagcgcctct ttcaccggcg cattcatact cgcaaattaa ccgctaatat 360
 cctcgcactt ggataatgtg tagccgacac ggaggagggg ggttgggggg gggttggggg 420
 gagacatgat ggtctgcccc acggatatta ttattttgtt gttttgtata attactgcgg 480
 caacattctc aaagggggccg tgccctcgcg cgggaaagcc catgacagag aattggacag 540
 ctccaagctc gcgatatact ctaacaacgg cgtgactcgg caatgaaggc ctgccgctcg 600
 agtgataggg cgaagtaaaa cggacgttac atgcggcact tagccggctg atgccggaga 660
 atacgggatt caacgataca atcacacgat gcgacacacc tcggcgactt ggcgctctat 720
 ggaagaaggc tgggttaaag ctggcgtaga ttttgcgcg cttggtttct taaccgggtt 780
 atttctatct ctcataatgcc gcgagcgaat gcgggggtgca gagcgcccgg gagtcgatgg 840
 tcctatcaga caagagcctg gccccggaac ctgggataat agaagccaaa ttaagccatg 900
 ggagtatcgt ccgggggtag gaaccgcacg ggcaactaga ggaggaagaa tttggtataa 960
 agggaggacg gcggaacagg cttgatggac atgaatcaga agacgacact gggcaactaa 1020
 acagcttgca gcagagtttt gtgccttgca taggccctcg atatcatggt ctcgttcact 1080
 ctcttcctca cggtcatcgc cgctgcggtg acgacggcca gccctctcga ggtgggtcaag 1140
 cgcgcatcc agccgggcac gggcaccac gaggggtact tctactcggt ctggaccgac 1200
 ggccgtggct cggtcgactt caaccccggg ccccgcggt cgtagacggt cacctggaac 1260

3123-4008PC.ST25.txt

aacgtcaaca actgggttgg cggcaagggc tggaacccgg gcccgcgcgc caagattgcg 1320
tacaacggca cctggaacaa ctacaacgtg aacagctgtg cgttgcctc ctctttctcc 1380
ctttcgcttg ttttccttga tgattgggat ccattttaaa agagaaggaa aaaaaaaca 1440
aagggaaata gaagataact aacgccaagc tctggcagac ctcgccctgt acggctggac 1500
tcgcaacccg ctggctgagt attacatcgt ggaggcatac ggcacgtaca acccctcgtc 1560
gggcacggcg cggctgggca ccatcgagga cgacggcggc gtgtacgaca tctacaagac 1620
gacgcggtac aaccagccgt ccatcgaggg gacctccacc ttcgaccagt actggtccgt 1680
ccgccgccag aagcgcgtcg gcggcactat cgacacgggc aagcactttg acgagtggaa 1740
gcgccagggc aacctccagc tcggcacctg gaactacatg atcatggcca ccgagggcta 1800
ccagagctct ggttcggcca ctatcgaggt ccgggaggcc taaagaagcc aggcgccttt 1860
cttttgtttt gcaggagggg gtagaggggg ggggggaggg aaaacgaaa gtagcagggg 1920
ggttttatgc cggcagccgt gggccattcg agtgcaacct gtatctctct ctctccaag 1980
tctccgggct ctttctcaga gaacttcaat atgtctgggg acaaaccacc ttgtgaaata 2040
caacggtaat tatctaagtt tgagtgccct atcgtatgct tctgaaaatt tcctgctcct 2100
tgatacaagt cggtttgagc cgagccaatg agactgtgtc gattgataga ggccctgaag 2160
gatcaagcgc gatgaacaa ttaagcatga ctacgtgcct agctgcagat aaatggaagc 2220
cactcaccaa ggtcaacccc gcatactggc acgtaagaac cttccgtgta caaggcccaa 2280
ccgactcaca tatctatctg cttgggtttt gggatgcggg tttttacca caaaacaaat 2340
ttgatacaat gctctgctgt gcccgggttg ctgagaccaa gccgtaatca gcgggcaggg 2400
aatcgagtag gtcacgcctg ttgcttggtc tagaacaac taatattaaa aagccttggtg 2460
ctcggcacac atacagaact cgacctgagg catgttcttg gaaggcggct agccagtcaa 2520
gtctggcacc aggcccttgg ctcgtcgagg ataccgaggg cgaggaggat gaggaagacc 2580
tctttctcgc ctcatatctc ttaggggacg aagaagacaa cgccggagcc acacaataat 2640
taggtctcat atcagacgtt tcggcctggc cgagctaata tgtctaatta tgcccatcag 2700
ccgtatgtcg aggcaggttg caccgatacg ctcgccgcgc cgcttcattc atctccgact 2760
gggcacaatg tcgccatctc ggccgtcaag gtggtgcaag atacctatta tgcaagcaga 2820
ggatcagatg gcgggccgat acgagcggct gctccggctt gcgagaaagc cgcttcgcag 2880
caaggatatc tggcaggccg ccattttcgg ttgggtattc tttgtcttgt ttgcttcgta 2940
attatgtcct ggctggcatt gtgggaaggg gcgaacctct tgatttccga tgggggtcga 3000

<210> 18
<211> 218
<212> PRT

3123-4008PC.ST25.txt

<213> Chrysosporium lucknowense

<400> 18

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

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

Gly Thr His Glu Gly Tyr Phe Tyr Ser Phe Trp Thr Asp Gly Arg Gly
 35 40 45

Ser Val Asp Phe Asn Pro Gly Pro Arg Gly Ser Tyr Ser Val Thr Trp
 50 55 60

Asn Asn Val Asn Asn Trp Val Gly Gly Lys Gly Trp Asn Pro Gly Pro
 65 70 75 80

Pro Arg Lys Ile Ala Tyr Asn Gly Thr Trp Asn Asn Tyr Asn Val Asn
 85 90 95

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

Tyr Ile Val Glu Ala Tyr Gly Thr Tyr Asn Pro Ser Ser Gly Thr Ala
 115 120 125

Arg Leu Gly Thr Ile Glu Asp Asp Gly Gly Val Tyr Asp Ile Tyr Lys
 130 135 140

Thr Thr Arg Tyr Asn Gln Pro Ser Ile Glu Gly Thr Ser Thr Phe Asp
 145 150 155 160

Gln Tyr Trp Ser Val Arg Arg Gln Lys Arg Val Gly Gly Thr Ile Asp
 165 170 175

Thr Gly Lys His Phe Asp Glu Trp Lys Arg Gln Gly Asn Leu Gln Leu
 180 185 190

Gly Thr Trp Asn Tyr Met Ile Met Ala Thr Glu Gly Tyr Gln Ser Ser
 195 200 205

Gly Ser Ala Thr Ile Glu Val Arg Glu Ala
 210 215