

(12) STANDARD PATENT
(19) AUSTRALIAN PATENT OFFICE

(11) Application No. **AU 2013323396 B2**

(54) Title
Production of fermentation products

(51) International Patent Classification(s)
C12P 7/16 (2006.01)

(21) Application No: **2013323396**

(22) Date of Filing: **2013.09.27**

(87) WIPO No: **WO14/052735**

(30) Priority Data

(31) Number
61/707,174

(32) Date
2012.09.28

(33) Country
US

(43) Publication Date: **2014.04.03**

(44) Accepted Journal Date: **2017.04.20**

(71) Applicant(s)
Butamax Advanced Biofuels LLC

(72) Inventor(s)
Dauner, Michael;Li, Sunny-Xiang;Burlew, Keith H.

(74) Agent / Attorney
Houlihan² Pty Ltd, Level 1 70 Doncaster Road, BALWYN NORTH, VIC, 3104, AU

(56) Related Art
US 2007/0092957 A1
WO 2011/063391 A1

(51) International Patent Classification:
C12P 7/16 (2006.01)(21) International Application Number:
PCT/US2013/062139(22) International Filing Date:
27 September 2013 (27.09.2013)

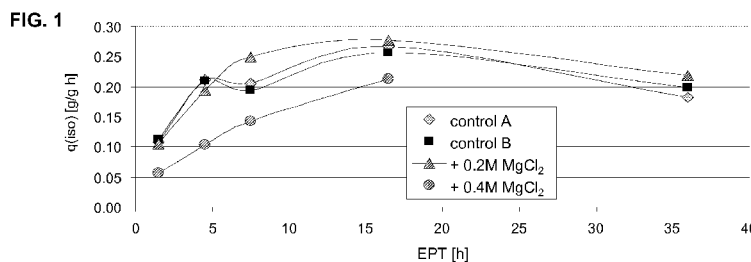
(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/707,174 28 September 2012 (28.09.2012) US(71) Applicant: BUTAMAX ADVANCED BIOFUELS LLC
[US/US]; Route 141 & Henry Clay, Dupont Experimental
Station, Building 356, Wilmington, Delaware 19880-0356
(US).(72) Inventors: DAUNER, Michael; 29 E. Avon Drive, Clay-
mont, Delaware 19703 (US). LI, Sunny-Xiang; 212 Cul-
len Way, Newark, Delaware 19711 (US). BURLEW,
Keith H.; 502 Burnham Lane, Middletown, Delaware
19709 (US).(74) Agents: PELLEGRINO, Susan M. et al.; E.I. du Pont de
Nemours and Company, Legal Patent Records Center,
Chestnut Run Plaza 721/2339, 974 Centre Road, Wilming-
ton, Delaware 19805 (US).(81) Designated States (unless otherwise indicated, for every
kind of national protection available): AE, AG, AL, AM,
AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY,BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM,
DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT,
HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KN, KP, KR,
KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME,
MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ,
OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA,
SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, 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, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ,
UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ,
TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK,
EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV,
MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM,
TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW,
KM, ML, MR, NE, SN, TD, TG).**Published:**

- with international search report (Art. 21(3))
- before the expiration of the time limit for amending the
claims and to be republished in the event of receipt of
amendments (Rule 48.2(h))
- with (an) indication(s) in relation to deposited biological
material furnished under Rule 13bis separately from the
description (Rules 13bis.4(d)(i) and 48.2(a)(viii))
- with sequence listing part of description (Rule 5.2(a))

(54) Title: PRODUCTION OF FERMENTATION PRODUCTS



(57) Abstract: The invention relates to processes for the production of fermentation products such as alcohols including ethanol and butanol, and the development of microorganisms capable of producing fermentation products via an engineered pathway in the microorganisms.

PRODUCTION OF FERMENTATION PRODUCTS

[0001] This application claims the benefit of U.S. Provisional Application No. 61/707,174, filed on September 28, 2012; the entire contents of which are herein incorporated by reference.

[0002] The Sequence Listing associated with this application is filed in electronic form via EFS-Web and hereby incorporated by reference into the specification in its entirety.

FIELD OF THE INVENTION

[0003] The invention relates to processes for the production of fermentation products such as alcohols including ethanol and butanol, and the development of microorganisms capable of producing fermentation products via an engineered pathway in the microorganisms.

BACKGROUND OF THE INVENTION

[0004] A number of chemicals and consumer products may be produced utilizing fermentation as the manufacturing process. For example, alcohols such as ethanol and butanol have a variety of industrial and scientific applications such as fuels, reagents, and solvents. Butanol is an important industrial chemical with a variety of applications including use as a fuel additive, as a feedstock chemical in the plastics industry, and as a food-grade extractant in the food and flavor industry. Each year 10 to 12 billion pounds of butanol are produced by chemical syntheses using starting materials derived from petrochemicals. The production of butanol or butanol isomers from materials such as plant-derived materials could minimize the use of petrochemicals and would represent an advance in the art. Furthermore, production of chemicals and fuels using plant-derived materials or other feedstock sources would provide eco-friendly and sustainable alternatives to petrochemical processes.

[0005] Techniques such as genetic engineering and metabolic engineering may be utilized to modify a microorganism to produce a certain product from plant-derived materials or other sources of feedstock. The microorganism may be modified, for example, by the insertion of genes such as the insertion of genes encoding a biosynthetic pathway, deletion of genes, or modifications to regulatory elements such as promoters. A microorganism may also be

engineered to improve cell productivity and yield, to eliminate by-products of biosynthetic pathways, and/or for strain improvement. Examples of microorganisms expressing engineered biosynthetic pathways for producing butanol isomers, including isobutanol, are described in U.S. Patent Nos. 7,851,188 and 7,993,889, the entire contents of each are herein incorporated by reference.

[0006] In order to develop an efficient and economical process for the production of butanol and other alcohols, productivity is an important factor. Productivity may be improved, for example, by increased growth of the microorganism, increased specific rates of glucose consumption and alcohol production, and increased yields and product titers. As such, the present invention is directed to the development of methods to improve productivity as well as the development of methods that produce fermentation products via an engineered pathway in the microorganisms.

SUMMARY OF THE INVENTION

[0007] The present invention is directed to a method for producing butanol comprising providing a recombinant host cell comprising a butanol biosynthetic pathway; and contacting the recombinant host cell with a fermentation medium comprising: a fermentable carbon substrate, magnesium, and backset wherein the backset is added to the fermentation medium as at least 5% of the water volume of the fermentation medium, wherein butanol is produced via the butanol biosynthetic pathway. In some embodiments, magnesium may be added to the fermentation medium. In some embodiments, magnesium may be added during propagation of the recombinant host cell. In some embodiments, magnesium or a portion thereof may be added as a magnesium salt or a concentrated magnesium salt solution. In some embodiments, magnesium in the fermentation medium may be in the range of about 5 mM to about 200 mM. In some embodiments, magnesium in the fermentation medium may be in the range of about 10 mM to about 150 mM. In some embodiments, magnesium in the fermentation medium may be in the range of about 30 mM to about 70 mM. In some embodiments, magnesium in the fermentation medium may be in the range of about 50 mM to about 150 mM. In some embodiments, the fermentation medium may comprise a low calcium-to-magnesium ratio or a high magnesium-to-calcium ratio. In some embodiments, magnesium may be added during preparation of the feedstock or biomass. In some embodiments, magnesium may be added during the fermentation process and/or during propagation of the recombinant host cell. In

some embodiments, the recombinant host cell may be pre-conditioned by the addition of magnesium.

[0008] The present invention is also directed to a method for producing butanol comprising providing a recombinant host cell comprising a butanol biosynthetic pathway; and contacting the recombinant host cell with a fermentation medium comprising: a fermentable carbon substrate and nutrients, wherein butanol is produced via the butanol biosynthetic pathway. In some embodiments, nutrients may be added to the fermentation medium. In some embodiments, nutrients may be added during propagation of the recombinant host cell. In some embodiments, nutrients may be added during preparation of feedstock. In some embodiments, nutrients may be added during the fermentation process and/or during propagation of the recombinant host cell. In some embodiments, the nutrients may comprise minerals, vitamins, amino acids, trace elements, other components, or mixtures thereof. In some embodiments, the nutrients may comprise one or more minerals, vitamins, amino acids, trace elements, and other components. In some embodiments, the nutrients may comprise calcium, iron, potassium, magnesium, manganese, sodium, phosphorus, sulfur, zinc, or mixtures thereof. In some embodiments, the nutrients may comprise one or more calcium, iron, potassium, magnesium, manganese, sodium, phosphorus, sulfur, and zinc. In some embodiments, the nutrients may be provided by the addition of backset. In some embodiments, backset may comprise minerals, vitamins, amino acids, trace elements, other components, or mixtures thereof. In some embodiments, backset may comprise one or more minerals, vitamins, amino acids, trace elements, other components. In some embodiments, backset may comprise minerals, vitamins, amino acids, calcium, iron, potassium, magnesium, manganese, sodium, phosphorus, sulfur, zinc, or mixtures thereof. In some embodiments, backset may comprise one or more minerals, vitamins, amino acids, calcium, iron, potassium, magnesium, manganese, sodium, phosphorus, sulfur, and zinc. In some embodiments, backset may comprise calcium, iron, potassium, magnesium, manganese, sodium, phosphorus, sulfur, zinc, or mixtures thereof. In some embodiments, backset may comprise one or more calcium, iron, potassium, magnesium, manganese, sodium, phosphorus, sulfur, and zinc.

[0009] In some embodiments, backset may be added to the feedstock, feedstock preparation, and/or fermentation medium. In some embodiments, backset is added to feedstock for the preparation of fermentation medium. In some embodiments, about 10% to

about 100% of backset (e.g., percentage of total backset generated by processing of whole stillage) may be added to feedstock, feedstock preparation, and/or fermentation medium. In some embodiments, about 10%, about 15%, about 20%, about 25%, about 30%, about 35%, about 40%, about 45%, about 50%, about 60%, about 70%, about 80%, about 90%, about 95%, or 100% of the backset may be added to feedstock, feedstock preparation, and/or fermentation medium. In some embodiments, backset may be added to feedstock, feedstock preparation, and/or fermentation medium as a percentage of the water volume of feedstock, feedstock preparation, and/or fermentation medium. In some embodiments, backset may be added as about 5%, about 10%, about 15%, about 20%, about 25%, about 30%, about 35%, about 40%, about 45%, or about 50% of the water volume of feedstock, feedstock preparation, and/or or fermentation medium.

[0010] In some embodiments, feedstock, feedstock preparation, and/or fermentation medium may be supplemented with backset. In some embodiments, backset is added to feedstock for the preparation of fermentation medium. In some embodiments, feedstock, feedstock preparation, and/or fermentation medium may be supplemented with about 10% to about 100% of backset (e.g., percentage of total backset generated by processing of whole stillage). In some embodiments, feedstock, feedstock preparation, and/or fermentation medium may be supplemented with about 10%, about 15%, about 20%, about 25%, about 30%, about 35%, about 40%, about 45%, about 50%, about 60%, about 70%, about 80%, about 90%, about 95%, or 100% of the backset. In some embodiments, feedstock, feedstock preparation, and/or fermentation medium may be supplemented with backset as a percentage of the water volume feedstock, feedstock preparation, and/or fermentation medium. In some embodiments, feedstock, feedstock preparation, and/or fermentation medium may be supplemented with backset as about 5%, about 10%, about 15%, about 20%, about 25%, about 30%, about 35%, about 40%, about 45%, or about 50% of the water volume of feedstock, feedstock preparation, and/or or fermentation medium.

[0011] In some embodiments, butanol may be 1-butanol, 2-butanol, 2-butanone, or isobutanol. In some embodiments, the butanol biosynthetic pathway may be an isobutanol biosynthetic pathway. In some embodiments, the isobutanol biosynthetic pathway may comprise a polynucleotide encoding a polypeptide that catalyzes a substrate to product conversion selected from the group consisting of: (a) pyruvate to acetolactate; (b) acetolactate to 2,3-dihydroxyisovalerate; (c) 2,3-dihydroxyisovalerate to 2-ketoisovalerate;

(d) 2-ketoisovalerate to isobutyraldehyde; and (e) isobutyraldehyde to isobutanol. In some embodiments, one or more of the substrate to product conversions may utilize reduced nicotinamide adenine dinucleotide (NADH) or reduced nicotinamide adenine dinucleotide phosphate (NADPH) as a cofactor. In some embodiments, NADH may be the preferred cofactor.

[0012] In some embodiments, the butanol biosynthetic pathway may comprise at least one polypeptide selected from the group having the following Enzyme Commission Numbers: EC 2.2.1.6, EC 1.1.1.86, EC 4.2.1.9, EC 4.1.1.72, EC 1.1.1.1, EC 1.1.1.265, EC 1.1.1.2, EC 1.2.4.4, EC 1.3.99.2, EC 1.2.1.57, EC 1.2.1.10, EC 2.6.1.66, EC 2.6.1.42, EC 1.4.1.9, EC 1.4.1.8, EC 4.1.1.14, EC 2.6.1.18, EC 2.3.1.9, EC 2.3.1.16, EC 1.1.1.30, EC 1.1.1.35, EC 1.1.1.157, EC 1.1.1.36, EC 4.2.1.17, EC 4.2.1.55, EC 1.3.1.44, EC 1.3.1.38, EC 5.4.99.13, EC 4.1.1.5, EC 2.7.1.29, EC 1.1.1.76, EC 1.2.1.57, and EC 4.2.1.28.

[0013] In some embodiments, the butanol biosynthetic pathway may comprise at least one polypeptide selected from the following group of enzymes: acetolactate synthase, acetohydroxy acid isomeroreductase, acetohydroxy acid dehydratase, branched-chain alpha-keto acid decarboxylase, branched-chain alcohol dehydrogenase, acylating aldehyde dehydrogenase, branched-chain keto acid dehydrogenase, butyryl-CoA dehydrogenase, butyraldehyde dehydrogenase, transaminase, valine dehydrogenase, valine decarboxylase, omega transaminase, acetyl-CoA acetyltransferase, 3-hydroxybutyryl-CoA dehydrogenase, crotonase, butyryl-CoA dehydrogenase, isobutyryl-CoA mutase, acetolactate decarboxylase, acetoin aminase, butanol dehydrogenase, butyraldehyde dehydrogenase, acetoin kinase, acetoin phosphate aminase, aminobutanol phosphate phospholyase, aminobutanol kinase, butanediol dehydrogenase, and butanediol dehydratase.

[0014] In some embodiments, the butanol biosynthetic pathway may comprise one or polynucleotides encoding polypeptides having acetolactate synthase, acetohydroxy acid isomeroreductase, acetohydroxy acid dehydratase, branched-chain alpha-keto acid decarboxylase, branched-chain alcohol dehydrogenase, acylating aldehyde dehydrogenase, branched-chain keto acid dehydrogenase, butyryl-CoA dehydrogenase, butyraldehyde dehydrogenase, transaminase, valine dehydrogenase, valine decarboxylase, omega transaminase, acetyl-CoA acetyltransferase, 3-hydroxybutyryl-CoA dehydrogenase, crotonase, butyryl-CoA dehydrogenase, isobutyryl-CoA mutase, acetolactate decarboxylase, acetoin aminase, butanol dehydrogenase, butyraldehyde dehydrogenase, acetoin kinase,

acetoin phosphate aminase, aminobutanol phosphate phospholyase, aminobutanol kinase, butanediol dehydrogenase, or butanediol dehydratase activity.

[0015] In some embodiments, the isobutanol biosynthetic pathway may comprise one or more polynucleotides encoding polypeptides having acetolactate synthase, keto acid reductoisomerase, dihydroxy acid dehydratase, ketoisovalerate decarboxylase, or alcohol dehydrogenase activity.

[0016] In some embodiments, the recombinant host cell may comprise a butanol biosynthetic pathway. In some embodiments, the butanol produced may be isobutanol. In some embodiments, the butanol produced may be 1-butanol. In some embodiments, the butanol produced may be 2-butanol. In some embodiments, the butanol produced may be 2-butanone.

[0017] In some embodiments, the microorganism may comprise an isobutanol biosynthetic pathway. In some embodiments, the microorganism may comprise a 1-butanol biosynthetic pathway. In some embodiments, the microorganism may comprise a 2-butanol biosynthetic pathway. In some embodiments, the microorganism may comprise a 2-butanone biosynthetic pathway.

[0018] In some embodiments, the recombinant host cell further may comprise a modification in a polynucleotide encoding a polypeptide having pyruvate decarboxylase activity. In some embodiments, the recombinant host cell may comprise a deletion, mutation, and/or substitution in an endogenous polynucleotide encoding a polypeptide having pyruvate decarboxylase activity. In some embodiments, the polypeptide having pyruvate decarboxylase activity may be selected from the group consisting of: PDC1, PDC5, PDC6, and combinations thereof. In some embodiments, the endogenous polynucleotide encoding a polypeptide having pyruvate decarboxylase activity may be selected from the group consisting of: PDC1, PDC5, PDC6, and combinations thereof. In some embodiments, the recombinant host cell may further comprise a deletion, mutation, and/or substitution in one or more endogenous polynucleotides encoding FRA2, GPD2, BDH1, and YMR.

[0019] In some embodiments, the recombinant host cell may be bacteria, cyanobacteria, filamentous fungi, or yeast. Suitable recombinant host cell capable of producing an alcohol via a biosynthetic pathway include a member of the genera *Clostridium*, *Zymomonas*, *Escherichia*, *Salmonella*, *Serratia*, *Erwinia*, *Klebsiella*, *Shigella*, *Rhodococcus*, *Pseudomonas*, *Bacillus*, *Lactobacillus*, *Enterococcus*, *Alcaligenes*, *Klebsiella*, *Paenibacillus*,

Arthrobacter, *Corynebacterium*, *Brevibacterium*, *Schizosaccharomyces*, *Kluyveromyces*, *Yarrowia*, *Pichia*, *Zygosaccharomyces*, *Debaryomyces*, *Candida*, *Brettanomyces*, *Pachysolen*, *Hansenula*, *Issatchenkia*, *Trichosporon*, *Yamadazyma*, or *Saccharomyces*. In some embodiments, the recombinant host cell may be selected from the group consisting of *Escherichia coli*, *Alcaligenes eutrophus*, *Bacillus licheniformis*, *Paenibacillus macerans*, *Rhodococcus erythropolis*, *Pseudomonas putida*, *Lactobacillus plantarum*, *Enterococcus faecium*, *Enterococcus gallinarum*, *Enterococcus faecalis*, *Bacillus subtilis*, *Candida sonorensis*, *Candida methanosorbosa*, *Kluyveromyces lactis*, *Kluyveromyces marxianus*, *Kluyveromyces thermotolerans*, *Issatchenkia orientalis*, *Debaryomyces hansenii*, and *Saccharomyces cerevisiae*. In some embodiments, the recombinant host cell may be yeast. In some embodiments, the recombinant host cell may be *Saccharomyces*, *Zygosaccharomyces*, *Schizosaccharomyces*, *Dekkera*, *Torulopsis*, *Brettanomyces*, and some species of *Candida*. In some embodiments, the recombinant host cell may be crabtree-positive yeast. Species of crabtree-positive yeast include, but are not limited to, *Saccharomyces cerevisiae*, *Saccharomyces kluyveri*, *Schizosaccharomyces pombe*, *Saccharomyces bayanus*, *Saccharomyces mikitaie*, *Saccharomyces paradoxus*, *Saccharomyces uvarum*, *Saccharomyces castelli*, *Saccharomyces kluyveri*, *Zygosaccharomyces rouxii*, *Zygosaccharomyces bailli*, and *Candida glabrata*.

[0020] The present invention is also directed to a composition comprising a recombinant host cell, a fermentable carbon substrate, magnesium and optionally alcohol, wherein the magnesium may be in the range of about 5 mM to about 200 mM. In some embodiments, magnesium may be in the range of about 10 mM to about 150 mM. In some embodiments, magnesium may be in the range of about 30 mM to about 70 mM. In some embodiments, magnesium may be in the range of about 50 mM to about 150 mM. In some embodiments, the composition may comprise a low calcium-to-magnesium ratio or a high magnesium-to-calcium ratio. In some embodiments, the alcohol is 1-butanol, 2-butanol, isobutanol, or 2-butanone.

[0021] The present invention is also directed to a composition comprising a recombinant host cell, a fermentable carbon substrate, nutrients, and optionally alcohol. In some embodiments, the recombinant host cell comprises a butanol biosynthetic pathway. In some embodiments, the butanol biosynthetic pathway is an isobutanol biosynthetic pathway. In some embodiments, the alcohol may be butanol. In some embodiments, the butanol may be

- 8 -

isobutanol. In some embodiments, the nutrients may comprise minerals, vitamins, amino acids, trace elements, other components, or mixtures thereof. In some embodiments, the nutrients may comprise calcium, iron, potassium, magnesium, manganese, sodium, phosphorus, sulfur, zinc, or mixtures thereof. In some embodiments, the composition may further comprise backset. In some embodiments, backset may comprise minerals, vitamins, amino acids, calcium, iron, potassium, magnesium, manganese, sodium, phosphorus, sulfur, zinc, or mixtures thereof. In some embodiments, backset may comprise calcium, iron, potassium, magnesium, manganese, sodium, phosphorus, sulfur, zinc, or mixtures thereof. In some embodiments, the composition may comprise backset in the amount of about 5%, about 10%, about 15%, about 20%, about 25%, about 30%, about 35%, about 40%, about 45%, or about 50% of the water volume of the composition.

[0022] The present invention is also directed to a composition comprising a recombinant host cell, a fermentable carbon substrate, backset, and optionally alcohol. In some embodiments, the recombinant host cell comprises a butanol biosynthetic pathway. In some embodiments, the butanol biosynthetic pathway is an isobutanol biosynthetic pathway. In some embodiments, the alcohol may be butanol. In some embodiments, the butanol may be isobutanol. In some embodiments, backset may comprise minerals, vitamins, amino acids, calcium, iron, potassium, magnesium, manganese, sodium, phosphorus, sulfur, zinc, or mixtures thereof. In some embodiments, backset may comprise calcium, iron, potassium, magnesium, manganese, sodium, phosphorus, sulfur, zinc, or mixtures thereof. In some embodiments, the composition may comprise backset in the amount of about 5%, about 10%, about 15%, about 20%, about 25%, about 30%, about 35%, about 40%, about 45%, or about 50% of the water volume of the composition.

[0023] The present invention is also directed to a composition comprising a recombinant host cell, a fermentable carbon substrate, and optionally alcohol. In some embodiments, the recombinant host cell comprises a butanol biosynthetic pathway. In some embodiments, the butanol biosynthetic pathway is an isobutanol biosynthetic pathway. In some embodiments, the composition may further comprise backset. In some embodiments, backset may comprise minerals, vitamins, amino acids, calcium, iron, potassium, magnesium, manganese, sodium, phosphorus, sulfur, zinc, or mixtures thereof. In some embodiments, backset may comprise calcium, iron, potassium, magnesium, manganese, sodium, phosphorus, sulfur, zinc, or mixtures thereof. In some embodiments, the composition may comprise backset in

the amount of about 5%, about 10%, about 15%, about 20%, about 25%, about 30%, about 35%, about 40%, about 45%, or about 50% of the water volume of the composition.

DESCRIPTION OF THE DRAWINGS

- [0024] The accompanying drawings, which are incorporated herein and form a part of the specification, illustrate the present invention and, together with the description, further serve to explain the principles of the invention and to enable a person skilled in the pertinent art to make and use the invention.
- [0025] Figure 1 shows average specific isobutanol production rates with and without magnesium supplementation (0.2 M and 0.4 M MgCl_2).
- [0026] Figure 2 demonstrates the formation of biomass with and without magnesium supplementation (0.05 M to 0.3 M MgCl_2).
- [0027] Figure 3 shows isobutanol concentrations in cultures with and without magnesium supplementation (0.05 M to 0.3 M MgCl_2).
- [0028] Figure 4 shows average specific isobutanol production rates with and without magnesium supplementation (0.05 M to 0.3 M MgCl_2).
- [0029] Figure 5 shows isobutanol concentrations in cultures supplemented with MgCl_2 or MgSO_4 .
- [0030] Figure 6 shows isobutanol concentrations in cultures supplemented with MgCl_2 or MgCl_2 and CaCl_2 .
- [0031] Figure 7 shows DHIV titers in cultures with and without magnesium supplementation.
- [0032] Figure 8 shows a concentration profile for isobutanol and DHIV in cultures with and without magnesium supplementation.
- [0033] Figure 9 shows isobutanol concentrations in cultures grown in corn mash medium with and without magnesium supplementation.
- [0034] Figure 10 shows isobutanol, glucose, and glycerol concentrations in cultures grown in corn mash medium with and without magnesium supplementation.
- [0035] Figures 11A-11D shows the effects of supplementation with backset on fermentation parameters with an isobutanologen.

[0036] Figures 12A-12D shows the effects of supplementation with backset on fermentation parameters with an ethanologen.

DESCRIPTION OF THE INVENTION

[0037] This invention is directed to processes for the production of fermentation products and to microorganisms that produce fermentation products and optimizations for producing fermentation products such as butanol at high rates and titers with advantaged economic process conditions.

[0038] With renewed interest in sustainable biofuels as an alternative energy source and the desire for the development of efficient and environmentally-friendly production methods, alcohol production using fermentation processes is a viable option to the current chemical synthesis processes. However, during fermentative production of alcohols, microorganisms may be subjected to various stress conditions including, for example, alcohol toxicity, oxidative stress, osmotic stress, and fluctuations in pH, temperature, and nutrient availability. The impact of these stress conditions can cause an inhibition of cell growth and decreased cell viability which can ultimately lead to a reduction in fermentation productivity and product yield. For example, some microorganisms that produce alcohol (e.g., ethanol, butanol) have low alcohol toxicity thresholds, and these low alcohol toxicity thresholds may limit the development of fermentation processes for the commercial production of alcohols. Thus, the ability to adjust fermentation conditions and/or metabolic processes to improve tolerance of the microorganism to stress conditions such as alcohol toxicity would be advantageous to maintain efficient alcohol production.

[0039] Magnesium is the most abundant divalent cation in cells, and predominantly serves as a counterion for solutes, for example, ATP and other nucleotides such as RNA and DNA. By binding to RNAs and many proteins, magnesium contributes to establishing and maintaining physiological structures. In addition, magnesium is an important cofactor in catalytic processes, for example, magnesium is a cofactor for enzymes such as glycolytic and fatty acid biosynthesis enzymes such as hexokinase, phosphofructokinase, phosphoglycerate kinase, enolase, and pyruvate kinase. Magnesium also has a role in membrane stability, cell metabolism, and cell growth and development. Calcium, a second messenger in signal transduction, regulates a number of cellular processes such as cell growth and cell division.

Calcium also has a role in maintenance of membrane permeability and stability, and regulation of lipid-protein interactions. As these cations are involved in various cellular functions, modification of the concentrations of magnesium and calcium in fermentation medium may have beneficial effects on cell viability and cell productivity. In addition, in some instances, calcium may have an inhibitory effect on magnesium-dependent enzymes. Thus, modifying concentrations of magnesium and calcium may have a beneficial effect on enzyme activity.

[0040] Stress conditions such as alcohol toxicity may lead to a disruption of cellular ionic homeostasis which can result in a reduction in cell growth, cell viability, and metabolic activity. Cations such as magnesium and calcium may remedy these detrimental effects by providing a protective effect. For example, magnesium appears to provide cellular protection against stress conditions such as ethanol toxicity and temperature (Dombek, et al., Appl. Environ. Microbiol. 52:975-981, 1986; Birch, et al. Enzyme Microb. Technol. 26:678-687, 2000. These protective effects of magnesium may result in improved alcohol production (e.g., rate and yield), glucose consumption, cell growth, and cell viability.

[0041] Magnesium, a cofactor for a number of enzymes, is required for the enzymatic activity of dihydroxyacid dehydratase (2,3-dihydroxy acid hydrolyase, E.C. 4.2.1.9) (see, e.g., Myers, J. Biol. Chem. 236:1414-1418, 1961; Xing, et al., J. Bacteriol. 173:2086-2092, 1991) and ketol-acid reductoisomerase (see, e.g., Chunduru, et al., Biochemistry 28:486-493, 1989; Tyagi, et al., FEBS Journal 272:593-602, 2005). Dihydroxyacid dehydratase catalyzes the conversion of 2,3-dihydroxyisovalerate to α -ketoisovalerate and ketol-acid reductoisomerase catalyzes the conversion (S)-acetolactate to 2,3-dihydroxyisovalerate, both steps in an isobutanol biosynthetic pathway. Adjustments to the concentrations of magnesium in fermentation medium may modify the enzymatic activity of dihydroxyacid dehydratase and ketol-acid reductoisomerase. For example, addition of magnesium may increase the enzymatic activity of dihydroxyacid dehydratase. Thus, supplementation of the fermentation medium with magnesium may improve the overall activity of a butanol biosynthetic pathway.

[0042] Fermentation medium may also be supplemented with other nutrients including, but not limited to, iron, zinc, and sulfur. Zinc is a cofactor for numerous enzymes such as peptidases, phospholipases, and enzymes involved in transcription, and structural proteins such as Zn finger proteins that regulate gene expression. Zinc also contributes to the

regulation of membrane fluidity. Iron, a redox protein cofactor, is required for the function of many metalloproteins such as catalases, hydrogenases, dehydrogenases, reductases, and acetyl-CoA synthases. In addition, iron may complex with sulfur to form iron-sulfur (Fe/S) clusters which serve as cofactors for various biological reactions including regulation of enzyme activity, mitochondrial respiration, ribosome biogenesis, cofactor biogenesis, gene expression regulation, and nucleotide metabolism. Supplementation of the fermentation medium with iron, zinc, and/or sulfur may also improve the overall activity of a butanol biosynthetic pathway.

[0043] The present invention is directed to methods of producing an alcohol by a fermentation process. In some embodiments, the method comprises cultivating a recombinant host cell as provided herein under conditions whereby the alcohol is produced and recovering the alcohol. In some embodiments, the alcohol may be butanol. In some embodiments, the alcohol may be 1-butanol, 2-butanol, 2-butanone, isobutanol, or tert-butanol. In some embodiments, the recombinant host cell may be contacted with a fermentation medium comprising: a fermentable carbon substrate and nutrients including, but not limited to, magnesium, calcium, zinc, iron, and sulfur. In some embodiments, one or more of the following; magnesium, calcium, zinc, iron, and sulfur may added to the fermentation medium.

[0044] In some embodiments, the recombinant host cell grown in supplemented fermentation medium exhibits increased alcohol production as compared to a recombinant host cell grown in non-supplemented fermentation medium. In some embodiments, alcohol production may be determined by measuring, for example: broth titer (grams alcohol produced per liter broth), alcohol yield (grams alcohol produced per gram substrate consumed or mol alcohol produced per mol substrate consumed), volumetric productivity (grams alcohol produced per liter per hour), specific productivity (grams alcohol produced per gram cell biomass per hour), or combinations thereof.

[0045] Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. In case of conflict, the present application including the definitions will control. Also, unless otherwise required by context, singular terms shall include pluralities and plural terms shall include the singular. All publications, patents and other references mentioned herein are incorporated by reference in their entireties for all purposes.

[0046] In order to further define this invention, the following terms and definitions are herein provided.

[0047] As used herein, the terms "comprises," "comprising," "includes," "including," "has," "having," "contains," or "containing," or any other variation thereof, will be understood to imply the inclusion of a stated integer or group of integers but not the exclusion of any other integer or group of integers. For example, a composition, a mixture, a process, a method, an article, or an apparatus that comprises a list of elements is not necessarily limited to only those elements but may include other elements not expressly listed or inherent to such composition, mixture, process, method, article, or apparatus. Further, unless expressly stated to the contrary, "or" refers to an inclusive or and not to an exclusive or. For example, a condition A or B is satisfied by any one of the following: A is true (or present) and B is false (or not present), A is false (or not present) and B is true (or present), and both A and B are true (or present).

[0048] As used herein, the term "consists of," or variations such as "consist of" or "consisting of," as used throughout the specification and claims, indicate the inclusion of any recited integer or group of integers, but that no additional integer or group of integers may be added to the specified method, structure, or composition.

[0049] As used herein, the term "consists essentially of," or variations such as "consist essentially of," or "consisting essentially of," as used throughout the specification and claims, indicate the inclusion of any recited integer or group of integers, and the optional inclusion of any recited integer or group of integers that do not materially change the basic or novel properties of the specified method, structure or composition. *See* M.P.E.P. § 2111.03.

[0050] Also, the indefinite articles "a" and "an" preceding an element or component of the invention are intended to be nonrestrictive regarding the number of instances, i.e., occurrences of the element or component. Therefore "a" or "an" should be read to include one or at least one, and the singular word form of the element or component also includes the plural unless the number is obviously meant to be singular.

[0051] The term "invention" or "present invention" as used herein is a non-limiting term and is not intended to refer to any single embodiment of the particular invention but encompasses all possible embodiments as described in the application.

[0052] As used herein, the term "about" modifying the quantity of an ingredient or reactant of the invention employed refers to variation in the numerical quantity that can occur, for

example, through typical measuring and liquid handling procedures used for making concentrates or solutions in the real world; through inadvertent error in these procedures; through differences in the manufacture, source, or purity of the ingredients employed to make the compositions or to carry out the methods; and the like. The term "about" also encompasses amounts that differ due to different equilibrium conditions for a composition resulting from a particular initial mixture. Whether or not modified by the term "about," the claims include equivalents to the quantities. In one embodiment, the term "about" means within 10% of the reported numerical value, or in some embodiments, within 5% of the reported numerical value.

[0053] The term "biomass" as used herein refers to the cell biomass of the fermentation product-producing microorganism, typically provided in units g/L dry cell weight (dcw).

[0054] The term "fermentation product" as used herein refers to any desired product of interest including lower alkyl alcohols such as butanol, lactic acid, 3-hydroxy-propionic acid, acrylic acid, acetic acid, succinic acid, citric acid, fumaric acid, malic acid, itaconic acid, 1,3-propane-diol, ethylene, glycerol, isobutyrate, etc.

[0055] The term "alcohol" as used herein refers to any alcohol that can be produced by a microorganism in a fermentation process. Alcohol includes any straight-chain or branched, saturated or unsaturated, alcohol molecule with 1-10 carbon atoms. For example, alcohol includes, but is not limited to, C₁ to C₈ alkyl alcohols. In some embodiments, alcohol is C₂ to C₈ alkyl alcohol. In other embodiments, the alcohol is C₂ to C₅ alkyl alcohol. It will be appreciated that C₁ to C₈ alkyl alcohols include, but are not limited to, methanol, ethanol, propanol, butanol, pentanol, and hexanol. Likewise, C₂ to C₈ alkyl alcohols include, but are not limited to, ethanol, propanol, butanol, pentanol, and hexanol. In some embodiments, alcohol may also include fusel alcohols (or fusel oils) and glycerol.

[0056] The term "butanol" or "butanol isomer" as used herein refers to 1-butanol, 2-butanol, 2-butanone, isobutanol, tert-butanol, or mixtures thereof. Isobutanol is also known as 2-methyl-1-propanol.

[0057] The term "butanol biosynthetic pathway" as used herein refers to an enzyme pathway to produce 1-butanol, 2-butanol, 2-butanone, or isobutanol. For example, butanol biosynthetic pathways are disclosed in U.S. Patent No. 7,993,889, the entire contents of which are herein incorporated by reference.

[0058] The term "isobutanol biosynthetic pathway" as used herein refers to an enzymatic pathway that produces isobutanol. From time to time "isobutanol biosynthetic pathway" is used synonymously with "isobutanol production pathway."

[0059] The term "2-butanone biosynthetic pathway" as used herein refers to an enzymatic pathway that produces 2-butanone.

[0060] The term "extractant" as used herein refers to one or more organic solvents which may be used to extract an alcohol from a fermentation broth.

[0061] A "recombinant host cell" as used herein refers to a host cell that has been genetically manipulated to express a biosynthetic production pathway, wherein the host cell either produces a biosynthetic product in greater quantities relative to an unmodified host cell or produces a biosynthetic product that is not ordinarily produced by an unmodified host cell. The term "recombinant host cell" and "recombinant microbial host cell" may be used interchangeably.

[0062] The term "engineered" as applied to a butanol biosynthetic pathway refers to the butanol biosynthetic pathway that is manipulated, such that the carbon flux from pyruvate through the engineered butanol biosynthetic pathway is maximized, thereby producing an increased amount of butanol directly from the fermentable carbon substrate. Such engineering includes expression of heterologous polynucleotides or polypeptides, overexpression of endogenous polynucleotides or polypeptides, cytosolic localization of proteins that do not naturally localize to cytosol, increased cofactor availability, decreased activity of competitive pathways, etc.

[0063] The term "butanologen" as used herein refers to a microorganism capable of producing butanol isomers. Such microorganisms may be recombinant host cells comprising an engineered butanol biosynthetic pathway. The term "isobutanologen" as used herein refers to a microorganism capable of producing isobutanol. Such microorganisms may be recombinant host cells comprising an engineered isobutanol biosynthetic pathway. The term "ethanologen" as used herein refers to a microorganism capable of producing ethanol. Such microorganisms may be recombinant host cells comprising an engineered ethanol biosynthetic pathway.

[0064] The term "fermentable carbon substrate" as used herein refers to a carbon source capable of being metabolized by microorganisms (or recombinant host cells) such as those disclosed herein. Suitable fermentable carbon substrates include, but are not limited to,

monosaccharides such as glucose or fructose; disaccharides such as lactose or sucrose; oligosaccharides; polysaccharides such as starch; cellulose; lignocellulose; hemicellulose; one-carbon substrates; fatty acids; and combinations thereof.

[0065] The term "fermentation medium" as used herein refers to a mixture of water, sugars (fermentable carbon substrates), dissolved solids, microorganisms producing fermentation products, fermentation product, and all other constituents of the material held in the fermentation vessel in which the fermentation product is being made by the reaction of fermentable carbon substrates to fermentation products, water and carbon dioxide (CO₂) by the microorganisms present. From time to time, as used herein the term "fermentation broth" and "fermentation mixture" can be used synonymously with "fermentation medium."

[0066] The term "feedstock" as used herein refers to a feed in a fermentation process, the feed containing a fermentable carbon source with or without undissolved solids and oil, and where applicable, the feed containing the fermentable carbon source before or after the fermentable carbon source has been removed from starch or obtained from the breakdown of complex sugars by further processing such as by liquefaction, saccharification, or other process. Suitable feedstocks include, but are not limited to, rye, wheat, corn, corn mash, cane, cane mash, barley, cellulosic material, lignocellulosic material, or mixtures thereof.

[0067] The term "magnesium salt" as used herein refers to non-solute ionic compounds containing the cation, magnesium. Examples of magnesium salt include, but are not limited to, magnesium chloride (MgCl₂) and magnesium sulfate (MgSO₄).

[0068] The term "concentrated magnesium salt solution" as used herein refers to solutions containing more than 100 mM dissolved magnesium.

[0069] The term "aerobic conditions" as used herein refers to growth conditions in the presence of oxygen.

[0070] The term "microaerobic conditions" as used herein refers to growth conditions with low levels of dissolved oxygen. For example, the oxygen level may be less than about 1% of air-saturation.

[0071] The term "anaerobic conditions" as used herein refers to growth conditions in the absence of oxygen.

[0072] The term "carbon substrate" as used herein refers to a carbon source capable of being metabolized by the microorganisms (or recombinant host cells) disclosed herein. Non-limiting examples of carbon substrates are provided herein and include, but are not limited

to, monosaccharides, oligosaccharides, polysaccharides, ethanol, lactate, succinate, glycerol, carbon dioxide, methanol, glucose, fructose, sucrose, xylose, arabinose, dextrose, and mixtures thereof.

[0073] The term "yield" as used herein refers to the amount of product per amount of carbon source in g/g. The yield may be exemplified for glucose as the carbon source. It is understood unless otherwise noted that yield is expressed as a percentage of the theoretical yield. In reference to a microorganism or metabolic pathway, "theoretical yield" is defined as the maximum amount of product that can be generated per total amount of substrate as dictated by the stoichiometry of the metabolic pathway used to make the product. It is understood that while in the present disclosure the yield is exemplified for glucose as a carbon source, the invention can be applied to other carbon sources and the yield may vary depending on the carbon source used. One skilled in the art can calculate yields on various carbon sources.

[0074] The term "titer" as used herein refers to the total amount of alcohol produced by fermentation per liter of fermentation medium. The total amount of alcohol includes: (i) the amount of alcohol in the fermentation medium; (ii) the amount of alcohol recovered from the organic extractant; and (iii) the amount of alcohol recovered from the gas phase, if gas stripping is used.

[0075] The term "rate" as used herein, refers to the total amount of alcohol produced by fermentation per liter of fermentation medium per hour of fermentation.

[0076] The term "growth rate" as used herein refers to the rate at which the microorganisms grow in the culture medium. The growth rate of the recombinant microorganisms can be monitored, for example, by measuring the optical density at 600 nanometers. The doubling time may be calculated from the logarithmic part of the growth curve and used as a measure of the growth rate.

Polypeptides and Polynucleotides for Use in the Invention

[0077] As used herein, the term "polypeptide" is intended to encompass a singular "polypeptide" as well as plural "polypeptides," and refers to a molecule composed of monomers (amino acids) linearly linked by amide bonds (also known as peptide bonds). The term "polypeptide" refers to any chain or chains of two or more amino acids, and does not refer to a specific length of the product. Thus, peptides, dipeptides, tripeptides,

oligopeptides, "protein," "amino acid chain," or any other term used to refer to a chain or chains of two or more amino acids, are included within the definition of "polypeptide," and the term "polypeptide" may be used instead of, or interchangeably with any of these terms. A polypeptide may be derived from a natural biological source or produced by recombinant technology, but is not necessarily translated from a designated nucleic acid sequence. It may be generated in any manner, including by chemical synthesis. The polypeptides used in this invention comprise full-length polypeptides and fragments thereof.

[0078] By an "isolated" polypeptide or a fragment, variant, or derivative thereof is intended a polypeptide that is not in its natural milieu. No particular level of purification is required. For example, an isolated polypeptide can be removed from its native or natural environment. Recombinantly produced polypeptides and proteins expressed in host cells are considered isolated for the purposes of the invention, as are native or recombinant polypeptides which have been separated, fractionated, or partially or substantially purified by any suitable technique.

[0079] A polypeptide of the invention may be of a size of about 10 or more, 20 or more, 25 or more, 50 or more, 75 or more, 100 or more, 200 or more, 500 or more, 1,000 or more, or 2,000 or more amino acids. Polypeptides may have a defined three-dimensional structure, although they do not necessarily have such structure. Polypeptides with a defined three-dimensional structure are referred to as folded, and polypeptides which do not possess a defined three-dimensional structure, but rather can adopt a large number of different conformations, and are referred to as unfolded.

[0080] Also included as polypeptides of the present invention are derivatives, analogs, or variants of the foregoing polypeptides, and any combination thereof. The terms "active variant," "active fragment," "active derivative," and "analog" refer to polypeptides of the present invention. Variants of polypeptides of the present invention include polypeptides with altered amino acid sequences due to amino acid substitutions, deletions, and/or insertions. Variants may occur naturally or be non-naturally occurring. Non-naturally occurring variants may be produced using art-known mutagenesis techniques. Variant polypeptides may comprise conservative or non-conservative amino acid substitutions, deletions and/or additions. Derivatives of polypeptides of the present invention, are polypeptides which have been altered so as to exhibit additional features not found on the native polypeptide. Examples include fusion proteins. Variant polypeptides may also be

referred to herein as "polypeptide analogs." As used herein, a "derivative" of a polypeptide refers to a polypeptide having one or more residues chemically derivatized by reaction of a functional side group. Also included as "derivatives" are those peptides which contain one or more naturally occurring amino acid derivatives of the twenty standard amino acids. For example, 4-hydroxyproline may be substituted for proline; 5-hydroxylysine may be substituted for lysine; 3-methylhistidine may be substituted for histidine; homoserine may be substituted for serine; and ornithine may be substituted for lysine.

[0081] A "fragment" is a unique portion of a polypeptide or other enzyme used in the invention which is identical in sequence to but shorter in length than the full-length parent sequence. A fragment may comprise up to the entire length of the defined sequence, minus one amino acid residue. For example, a fragment may comprise from 5 to 1000 contiguous amino acid residues. A fragment may be at least 5, 10, 15, 16, 20, 25, 30, 40, 50, 60, 75, 100, 150, 250 or at least 500 contiguous amino acid residues in length. Fragments may be preferentially selected from certain regions of a molecule. For example, a polypeptide fragment may comprise a certain length of contiguous amino acids selected from the first 100 or 200 amino acids of a polypeptide as shown in a certain defined sequence. Clearly, these lengths are exemplary, and any length that is supported by the specification, including the Sequence Listing, may be encompassed by the present embodiments.

[0082] Alternatively, recombinant variants encoding these same or similar polypeptides can be synthesized or selected by making use of the "redundancy" in the genetic code. Various codon substitutions, such as the silent changes which produce various restriction sites, may be introduced to optimize cloning into a plasmid or viral vector or expression in a host cell system.

[0083] Amino acid "substitutions" may be the result of replacing one amino acid with another amino acid having similar structural and/or chemical properties, i.e., conservative amino acid replacements, or they can be the result of replacing one amino acid with an amino acid having different structural and/or chemical properties, i.e., non-conservative amino acid replacements. "Conservative" amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues involved. For example, nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan, and methionine; polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine,

asparagine, and glutamine; positively charged (basic) amino acids include arginine, lysine, and histidine; and negatively charged (acidic) amino acids include aspartic acid and glutamic acid. Alternatively, "non-conservative" amino acid substitutions can be made by selecting the differences in polarity, charge, solubility, hydrophobicity, hydrophilicity, or the amphipathic nature of any of these amino acids. "Insertions" or "deletions" may be in the range of about 1 to about 20 amino acids, or may be in the range of about 1 to 10 amino acids. The variation allowed may be experimentally determined by systematically making insertions, deletions, or substitutions of amino acids in a polypeptide molecule using recombinant DNA techniques and assaying the resulting recombinant variants for activity.

[0084] As used herein, the term "variant" refers to a polypeptide differing from a specifically recited polypeptide of the invention by amino acid insertions, deletions, mutations, and substitutions, created using, for example, recombinant DNA techniques, such as mutagenesis. Guidance in determining which amino acid residues may be replaced, added, or deleted without abolishing activities of interest, may be found by comparing the sequence of the particular polypeptide with that of homologous polypeptides, for example, yeast or bacterial, and minimizing the number of amino acid sequence changes made in regions of high homology (conserved regions) or by replacing amino acids with consensus sequences.

[0085] By a polypeptide having an amino acid or polypeptide sequence at least, for example, 95% "identical" to a query amino acid sequence of the present invention, it is intended that the amino acid sequence of the subject polypeptide is identical to the query sequence except that the subject polypeptide sequence may include up to five amino acid alterations per each 100 amino acids of the query amino acid sequence. In other words, to obtain a polypeptide having an amino acid sequence at least 95% identical to a query amino acid sequence, up to 5% of the amino acid residues in the subject sequence may be inserted, deleted, or substituted with another amino acid. These alterations of the reference sequence may occur at the amino or carboxy terminal positions of the reference amino acid sequence or anywhere between those terminal positions, interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence.

[0086] As a practical matter, whether any particular polypeptide is at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identical to a reference polypeptide can be determined

conventionally using known computer programs. One method for determining the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, is using the FASTDB computer program based on the algorithm of Brutlag, et al. (Comp. Appl. Biosci. 6:237-245, 1990). In a sequence alignment, the query and subject sequences are either both nucleotide sequences or both amino acid sequences. The result of the global sequence alignment is in percent identity. Example parameters used in a FASTDB amino acid alignment are: Matrix=PAM 0, k-tuple=2, Mismatch Penalty=1, Joining Penalty=20, Randomization Group Length=0, Cutoff Score=1, Window Size=sequence length, Gap Penalty=5, Gap Size Penalty=0.05, Window Size=500 or the length of the subject amino acid sequence, whichever is shorter.

[0087] If the subject sequence is shorter than the query sequence due to N- or C-terminal deletions, not because of internal deletions, a manual correction must be made to the results. This is because the FASTDB program does not account for N- and C-terminal truncations of the subject sequence when calculating global percent identity. For subject sequences truncated at the N- and C-termini, relative to the query sequence, the percent identity is corrected by calculating the number of residues of the query sequence that are N- and C-terminal of the subject sequence, which are not matched/aligned with a corresponding subject residue, as a percent of the total bases of the query sequence. Whether a residue is matched/aligned is determined by results of the FASTDB sequence alignment. This percentage is then subtracted from the percent identity, calculated by the FASTDB program using the specified parameters, to arrive at a final percent identity score. This final percent identity score is what is used for the purposes of the present invention. Only residues to the N- and C-termini of the subject sequence, which are not matched/aligned with the query sequence, are considered for the purposes of manually adjusting the percent identity score. That is, only query residue positions outside the farthest N- and C-terminal residues of the subject sequence.

[0088] For example, a 90 amino acid residue subject sequence is aligned with a 100 residue query sequence to determine percent identity. The deletion occurs at the N-terminus of the subject sequence and therefore, the FASTDB alignment does not show a matching/alignment of the first 10 residues at the N-terminus. The 10 unpaired residues represent 10% of the sequence (number of residues at the N- and C-termini not matched/total number of residues in the query sequence) so 10% is subtracted from the percent identity

- 22 -

score calculated by the FASTDB program. If the remaining 90 residues were perfectly matched the final percent identity would be 90%. In another example, a 90 residue subject sequence is compared with a 100 residue query sequence. This time the deletions are internal deletions so there are no residues at the N- or C-termini of the subject sequence which are not matched/aligned with the query. In this case, the percent identity calculated by FASTDB is not manually corrected. Once again, only residue positions outside the N- and C-terminal ends of the subject sequence, as displayed in the FASTDB alignment, which are not matched/aligned with the query sequence are manually corrected for. No other manual corrections are to be made for the purposes of the present invention.

[0089] Polypeptides and other enzymes suitable for use in the present invention and fragments thereof are encoded by polynucleotides. The term "polynucleotide" is intended to encompass a singular nucleic acid as well as plural nucleic acids, and refers to an isolated nucleic acid molecule or construct, for example, messenger RNA (mRNA), virally-derived RNA, or plasmid DNA (pDNA). A polynucleotide may comprise a conventional phosphodiester bond or a non-conventional bond (e.g., an amide bond, such as found in peptide nucleic acids (PNA)). A polynucleotide can contain the nucleotide sequence of the full-length cDNA sequence, or a fragment thereof, including the untranslated 5' and 3' sequences and the coding sequences. The polynucleotide can be composed of any polyribonucleotide or polydeoxyribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. For example, polynucleotides can be composed of single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. "Polynucleotide" embraces chemically, enzymatically, or metabolically modified forms.

[0090] The term "nucleic acid" refers to any one or more nucleic acid segments, for example, DNA or RNA fragments, present in a polynucleotide. Polynucleotides according to the present invention further include such molecules produced synthetically. Polynucleotides of the invention may be native to the host cell or heterologous. In addition, a polynucleotide or a nucleic acid may be or may include a regulatory element such as a promoter, ribosome binding site, or a transcription terminator.

[0091] In certain embodiments, the polynucleotide or nucleic acid is DNA. In the case of DNA, a polynucleotide comprising a nucleic acid, which encodes a polypeptide normally may include a promoter and/or other transcription or translation control elements operably associated with one or more coding regions. An operable association is when a coding region for a gene product, for example, a polypeptide, is associated with one or more regulatory sequences in such a way as to place expression of the gene product under the influence or control of the regulatory sequence(s). Two DNA fragments (such as a polypeptide coding region and a promoter associated therewith) are "operably associated" if induction of promoter function results in the transcription of mRNA encoding the desired gene product and if the nature of the linkage between the two DNA fragments does not interfere with the ability of the expression regulatory sequences to direct the expression of the gene product or interfere with the ability of the DNA template to be transcribed. Thus, a promoter region would be operably associated with a nucleic acid encoding a polypeptide if the promoter was capable of effecting transcription of that nucleic acid. Other transcription control elements include, for example, enhancers, operators, repressors, and transcription termination signals, which can be operably associated with the polynucleotide. Promoters and other transcription control regions are known to those of skill in the art.

[0092] A polynucleotide sequence can be referred to as "isolated," if it has been removed from its native environment. For example, a heterologous polynucleotide encoding a polypeptide or polypeptide fragment having enzymatic activity (e.g., the ability to convert a substrate to product) contained in a vector is considered isolated for the purposes of the present invention. Further examples of an isolated polynucleotide include recombinant polynucleotides maintained in heterologous host cells or purified (partially or substantially) polynucleotides in solution. Isolated polynucleotides or nucleic acids according to the present invention further include such molecules produced synthetically. An isolated polynucleotide fragment in the form of a polymer of DNA can be comprised of one or more segments of cDNA, genomic DNA, or synthetic DNA.

[0093] The term "gene" refers to a nucleic acid fragment that is capable of being expressed as a specific protein, optionally including regulatory sequences preceding (5' non-coding sequences) and following (3' non-coding sequences) the coding sequence.

[0094] As used herein, a "coding region" or "ORF" is a portion of nucleic acid which consists of codons translated into amino acids. Although a "stop codon" (TAG, TGA, or

TAA) is not translated into an amino acid, it may be considered to be part of a coding region, if present, but any flanking sequences, for example, promoters, ribosome binding sites, transcriptional terminators, introns, 5' and 3' non-translated regions, and the like, are not part of a coding region. "Regulatory sequences" refer to nucleotide sequences located upstream (5' non-coding sequences), within, or downstream (3' non-coding sequences) of a coding sequence that influence the transcription, RNA processing or stability, or translation of the associated coding sequence. Regulatory sequences can include promoters, translation leader sequences, introns, polyadenylation recognition sequences, RNA processing sites, effector binding sites, and stem-loop structures.

[0095] A variety of translation control elements are known to those of ordinary skill in the art. These include, but are not limited to, ribosome binding sites, translation initiation and termination codons, and elements derived from viral systems (particularly an internal ribosome entry site, or IRES). In other embodiments, a polynucleotide of the present invention is RNA, for example, in the form of messenger RNA (mRNA). RNA of the present invention may be single-stranded or double-stranded.

[0096] Polynucleotide and nucleic acid coding regions of the present invention may be associated with additional coding regions which encode secretory or signal peptides, which direct the secretion of a polypeptide encoded by a polynucleotide of the present invention.

[0097] As used herein, the term "transformation" refers to the transfer of a nucleic acid fragment into the genome of a host organism, resulting in genetically stable inheritance. Host organisms containing the transformed nucleic acid fragments are referred to as "recombinant" or "transformed" organisms.

[0098] The term "expression," as used herein refers to the transcription and stable accumulation of sense (mRNA) or antisense RNA derived from the nucleic acid fragment of the invention. Expression may also refer to translation of mRNA into a polypeptide.

[0099] The term "overexpression," as used herein, refers to an increase in the level of nucleic acid or protein in a host cell. Thus, overexpression can result from increasing the level of transcription or translation of an endogenous sequence in a host cell or can result from the introduction of a heterologous sequence into a host cell. Overexpression can also result from increasing the stability of a nucleic acid or protein sequence.

[00100] The terms "plasmid," "vector," and "cassette" refer to an extra chromosomal element often carrying genes which are not part of the central metabolism of the cell, and

usually in the form of circular double-stranded DNA fragments. Such elements may be autonomously replicating sequences, genome integrating sequences, phage or nucleotide sequences, linear or circular, of a single- or double-stranded DNA or RNA, derived from any source, in which a number of nucleotide sequences have been joined or recombined into a unique construction which is capable of introducing a promoter fragment and DNA sequence for a selected gene product along with appropriate 3' untranslated sequence into a cell. "Transformation cassette" refers to a specific vector containing a foreign gene and having elements in addition to the foreign gene that facilitates transformation of a particular host cell. "Expression cassette" refers to a specific vector containing a foreign gene and having elements in addition to the foreign gene that allow for enhanced expression of that gene in a foreign host.

[00101] The term "artificial" refers to a synthetic, or non-host cell derived composition, for example, a chemically-synthesized oligonucleotide.

[00102] As used herein, "native" refers to the form of a polynucleotide, gene, or polypeptide as found in nature with its own regulatory sequences, if present.

[00103] The term "endogenous" when used in reference to a polynucleotide, a gene, or a polypeptide refers to a native polynucleotide or gene in its natural location in the genome of an organism, or for a native polypeptide, is transcribed and translated from this location in the genome.

[00104] The term "heterologous" when used in reference to a polynucleotide, a gene, or a polypeptide refers to a polynucleotide, gene, or polypeptide not normally found in the host organism. "Heterologous polynucleotide" includes a native coding region, or portion thereof, that is reintroduced into the source organism in a form that is different from the corresponding native polynucleotide. "Heterologous gene" includes a native coding region, or portion thereof, that is reintroduced into the source organism in a form that is different from the corresponding native gene, for example, not in its natural location in the organism's genome. For example, a heterologous gene may include a native coding region that is a portion of a chimeric gene including non-native regulatory regions that is reintroduced into the native host. A "transgene" is a gene that has been introduced into the genome by a transformation procedure. "Heterologous polypeptide" includes a native polypeptide that is reintroduced into the source organism in a form that is different from the corresponding

native polypeptide. The heterologous polynucleotide or gene may be introduced into the host organism, for example, by gene transfer.

[00105] As used herein, the term "modification" refers to a change in a polynucleotide disclosed herein that results in altered activity of a polypeptide encoded by the polynucleotide, as well as a change in a polypeptide disclosed herein that results in altered activity of the polypeptide. Such changes can be made by methods well known in the art, including, but not limited to, deleting, mutating (e.g., spontaneous mutagenesis, random mutagenesis, mutagenesis caused by mutator genes, or transposon mutagenesis), substituting, inserting, altering the cellular location, altering the state of the polynucleotide or polypeptide (e.g., methylation, phosphorylation, or ubiquitination), removing a cofactor, chemical modification, covalent modification, irradiation with UV or X-rays, homologous recombination, mitotic recombination, promoter replacement methods, and/or combinations thereof. Guidance in determining which nucleotides or amino acid residues can be modified, can be found by comparing the sequence of the particular polynucleotide or polypeptide with that of homologous polynucleotides or polypeptides, for example, yeast or bacterial, and maximizing the number of modifications made in regions of high homology (conserved regions) or consensus sequences.

[00106] As used herein, the term "variant" refers to a polynucleotide differing from a specifically recited polynucleotide of the invention by nucleotide insertions, deletions, mutations, and substitutions, created using, for example, recombinant DNA techniques, such as mutagenesis. Recombinant polynucleotide variants encoding same or similar polypeptides may be synthesized or selected by making use of the "redundancy" in the genetic code. Various codon substitutions, such as silent changes which produce various restriction sites, may be introduced to optimize cloning into a plasmid or viral vector for expression. Mutations in the polynucleotide sequence may be reflected in the polypeptide or domains of other peptides added to the polypeptide to modify the properties of any part of the polypeptide.

[00107] The term "recombinant genetic expression element" refers to a nucleic acid fragment that expresses one or more specific proteins, including regulatory sequences preceding (5' non-coding sequences) and following (3' termination sequences) coding sequences for the proteins. A chimeric gene is a recombinant genetic expression element.

The coding regions of an operon may form a recombinant genetic expression element, along with an operably linked promoter and termination region.

[00108] "Regulatory sequences" refers to nucleotide sequences located upstream (5' non-coding sequences), within, or downstream (3' non-coding sequences) of a coding sequence, and which influence the transcription, RNA processing or stability, or translation of the associated coding sequence. Regulatory sequences may include promoters, enhancers, operators, repressors, transcription termination signals, translation leader sequences, introns, polyadenylation recognition sequences, RNA processing site, effector binding site and stem-loop structure.

[00109] The term "promoter" refers to a nucleic acid sequence capable of controlling the expression of a coding sequence or functional RNA. In general, a coding sequence is located 3' to a promoter sequence. Promoters may be derived in their entirety from a native gene, or be composed of different elements derived from different promoters found in nature, or even comprise synthetic nucleic acid segments. It is understood by those skilled in the art that different promoters may direct the expression of a gene in different tissues or cell types, or at different stages of development, or in response to different environmental or physiological conditions. Promoters which cause a gene to be expressed in most cell types at most times are commonly referred to as "constitutive promoters." "Inducible promoters," on the other hand, cause a gene to be expressed when the promoter is induced or turned on by a promoter-specific signal or molecule. It is further recognized that since in most cases the exact boundaries of regulatory sequences have not been completely defined, DNA fragments of different lengths may have identical promoter activity. For example, it will be understood that "FBA1 promoter" can be used to refer to a fragment derived from the promoter region of the FBA1 gene.

[00110] The term "terminator" as used herein refers to DNA sequences located downstream of a coding sequence. This includes polyadenylation recognition sequences and other sequences encoding regulatory signals capable of affecting mRNA processing or gene expression. The polyadenylation signal is usually characterized by affecting the addition of polyadenylic acid tracts to the 3' end of the mRNA precursor. The 3' region can influence the transcription, RNA processing or stability, or translation of the associated coding sequence. It is recognized that since in most cases the exact boundaries of regulatory sequences have not been completely defined, DNA fragments of different lengths may have

identical terminator activity. For example, it will be understood that "CYC1 terminator" can be used to refer to a fragment derived from the terminator region of the CYC1 gene.

[00111] The term "operably linked" refers to the association of nucleic acid sequences on a single nucleic acid fragment so that the function of one is affected by the other. For example, a promoter is operably linked with a coding sequence when it is capable of effecting the expression of that coding sequence (i.e., that the coding sequence is under the transcriptional control of the promoter). Coding sequences can be operably linked to regulatory sequences in sense or antisense orientation.

[00112] As used herein the term "transformation" refers to the transfer of a nucleic acid fragment into the genome of a host microorganism, resulting in genetically stable inheritance. Host microorganisms containing the transformed nucleic acid fragments are referred to as "transgenic," "recombinant" or "transformed" microorganisms.

[00113] The term "codon-optimized" as it refers to genes or coding regions of nucleic acid molecules for transformation of various hosts, refers to the alteration of codons in the gene or coding regions of the nucleic acid molecules to reflect the typical codon usage of the host organism without altering the polypeptide encoded by the DNA. Such optimization includes replacing at least one, or more than one, or a significant number, of codons with one or more codons that are more frequently used in the genes of that organism.

[00114] Deviations in the nucleotide sequence that comprise the codons encoding the amino acids of any polypeptide chain allow for variations in the sequence coding for the gene. Since each codon consists of three nucleotides, and the nucleotides comprising DNA are restricted to four specific bases, there are 64 possible combinations of nucleotides, 61 of which encode amino acids (the remaining three codons encode signals ending translation). The "genetic code" which shows which codons encode which amino acids is reproduced herein as Table 1. As a result, many amino acids are designated by more than one codon. For example, the amino acids alanine and proline are coded for by four triplets, serine and arginine by six, whereas tryptophan and methionine are coded by just one triplet. This degeneracy allows for DNA base composition to vary over a wide range without altering the amino acid sequence of the proteins encoded by the DNA.

Table 1: The Standard Genetic Code

	T	C	A	G
T	TTT Phe (F) TTC " TTA Leu (L) TTG "	TCT Ser (S) TCC " TCA " TCG "	TAT Tyr (Y) TAC " TAA Ter TAG Ter	TGT Cys (C) TGC TGA Ter TGG Trp (W)
C	CTT Leu (L) CTC " CTA " CTG "	CCT Pro (P) CCC " CCA " CCG "	CAT His (H) CAC " CAA Gln (Q) CAG "	CGT Arg (R) CGC " CGA " CGG "
A	ATT Ile (I) ATC " ATA " ATG Met (M)	ACT Thr (T) ACC " ACA " ACG "	AAT Asn (N) AAC " AAA Lys (K) AAG "	AGT Ser (S) AGC " AGA Arg (R) AGG "
G	GTT Val (V) GTC " GTA " GTG "	GCT Ala (A) GCC " GCA " GCG "	GAT Asp (D) GAC " GAA Glu (E) GAG "	GGT Gly (G) GGC " GGA " GGG "

[00115] Many organisms display a bias for use of particular codons to code for insertion of a particular amino acid in a growing peptide chain. Codon preference or codon bias, differences in codon usage between organisms, is afforded by degeneracy of the genetic code, and is well documented among many organisms. Codon bias often correlates with the efficiency of translation of messenger RNA (mRNA), which is in turn believed to be dependent on, *inter alia*, the properties of the codons being translated and the availability of particular transfer RNA (tRNA) molecules. The predominance of selected tRNAs in a cell is generally a reflection of the codons used most frequently in peptide synthesis. Accordingly, genes can be tailored for optimal gene expression in a given organism based on codon optimization.

[00116] Given the large number of gene sequences available for a wide variety of animal, plant, and microbial species, it is possible to calculate the relative frequencies of codon usage. Codon usage tables are readily available, for example, at the "Codon Usage Database" available at <http://www.kazusa.or.jp/codon/> (visited March 20, 2008), and these tables can be adapted in a number of ways (see, e.g., Nakamura, et al., Nucl. Acids Res. 28:292, 2000). Codon usage tables for yeast, calculated from GenBank Release 128.0 [15

February 2002], are reproduced below as Table 2. This table uses mRNA nomenclature, and so instead of thymine (T) which is found in DNA, the tables use uracil (U) which is found in RNA. The Table has been adapted so that frequencies are calculated for each amino acid, rather than for all 64 codons.

Table 2: Codon Usage Table for *Saccharomyces cerevisiae* Genes

Amino Acid	Codon	Number	Frequency per thousand
Phe	UUU	170666	26.1
Phe	UUC	120510	18.4
Leu	UUA	170884	26.2
Leu	UUG	177573	27.2
Leu	CUU	80076	12.3
Leu	CUC	35545	5.4
Leu	CUA	87619	13.4
Leu	CUG	68494	10.5
Ile	AUU	196893	30.1
Ile	AUC	112176	17.2
Ile	AUA	116254	17.8
Met	AUG	136805	20.9
Val	GUU	144243	22.1
Val	GUC	76947	11.8
Val	GUA	76927	11.8
Val	GUG	70337	10.8
Ser	UCU	153557	23.5
Ser	UCC	92923	14.2
Ser	UCA	122028	18.7
Ser	UCG	55951	8.6
Ser	AGU	92466	14.2
Ser	AGC	63726	9.8
Pro	CCU	88263	13.5
Pro	CCC	44309	6.8
Pro	CCA	119641	18.3
Pro	CCG	34597	5.3
Thr	ACU	132522	20.3
Thr	ACC	83207	12.7

- 31 -

Amino Acid	Codon	Number	Frequency per thousand
Thr	ACA	116084	17.8
Thr	ACG	52045	8.0
Ala	GCU	138358	21.2
Ala	GCC	82357	12.6
Ala	GCA	105910	16.2
Ala	GCG	40358	6.2
Tyr	UAU	122728	18.8
Tyr	UAC	96596	14.8
His	CAU	89007	13.6
His	CAC	50785	7.8
Gln	CAA	178251	27.3
Gln	CAG	79121	12.1
Asn	AAU	233124	35.7
Asn	AAC	162199	24.8
Lys	AAA	273618	41.9
Lys	AAG	201361	30.8
Asp	GAU	245641	37.6
Asp	GAC	132048	20.2
Glu	GAA	297944	45.6
Glu	GAG	125717	19.2
Cys	UGU	52903	8.1
Cys	UGC	31095	4.8
Trp	UGG	67789	10.4
Arg	CGU	41791	6.4
Arg	CGC	16993	2.6
Arg	CGA	19562	3.0
Arg	CGG	11351	1.7
Arg	AGA	139081	21.3
Arg	AGG	60289	9.2
Gly	GGU	156109	23.9
Gly	GGC	63903	9.8

- 32 -

Amino Acid	Codon	Number	Frequency per thousand
Gly	GGA	71216	10.9
Gly	GGG	39359	6.0
Stop	UAA	6913	1.1
Stop	UAG	3312	0.5
Stop	UGA	4447	0.7

[00117] By utilizing this or similar tables, one of ordinary skill in the art can apply the frequencies to any given polypeptide sequence, and produce a nucleic acid fragment of a codon-optimized coding region which encodes the polypeptide, but which uses codons optimal for a given species.

[00118] Randomly assigning codons at an optimized frequency to encode a given polypeptide sequence can be done manually by calculating codon frequencies for each amino acid, and then assigning the codons to the polypeptide sequence randomly. Additionally, various algorithms and computer software programs are readily available to those of ordinary skill in the art. For example, the "EditSeq" function in the Lasergene® Package (DNASTAR, Inc., Madison, WI), the backtranslation function in the VectorNTI Suite (InforMax, Inc., Bethesda, MD), and the backtranslate function in the GCG--Wisconsin Package (Accelrys, Inc., San Diego, CA). In addition, various resources are publicly available to codon-optimize coding region sequences, for example, the backtranslation function at <http://www.entelechon.com/bioinformatics/backtranslation.php?lang=eng> (visited April 15, 2008) and the backtranseq function available at <http://bioinfo.pbi.nrc.ca:8090/EMBOSS/index.html> (visited July 9, 2002). Constructing a rudimentary algorithm to assign codons based on a given frequency can also easily be accomplished with basic mathematical functions by one of ordinary skill in the art. Codon-optimized coding regions can be designed by various methods known to those skilled in the art including software packages such as "synthetic gene designer" (<http://phenotype.biosci.umbc.edu/codon/sgd/index.php>).

[00119] A polynucleotide or nucleic acid fragment is "hybridizable" to another nucleic acid fragment, such as a cDNA, genomic DNA, or RNA molecule, when a single-stranded form of the nucleic acid fragment can anneal to the other nucleic acid fragment under the appropriate conditions of temperature and solution ionic strength. Hybridization and washing conditions are well known and exemplified, for example, in Sambrook, J., Fritsch,

E. F. and Maniatis, T. *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory: Cold Spring Harbor, NY (1989), particularly Chapter 11 and Table 11.1 therein. The conditions of temperature and ionic strength determine the "stringency" of the hybridization. Stringency conditions can be adjusted to screen for moderately similar fragments (such as homologous sequences from distantly related organisms), to highly similar fragments (such as genes that duplicate functional enzymes from closely related organisms). Post-hybridization washes determine stringency conditions. One set of preferred conditions uses a series of washes starting with 6X SSC, 0.5% SDS at room temperature for 15 min, then repeated with 2X SSC, 0.5% SDS at 45°C for 30 min, and then repeated twice with 0.2X SSC, 0.5% SDS at 50°C for 30 min. A more preferred set of stringent conditions uses higher temperatures in which the washes are identical to those above except for the temperature of the final two 30 min washes in 0.2X SSC, 0.5% SDS was increased to 60°C. Another preferred set of highly stringent conditions uses two final washes in 0.1X SSC, 0.1% SDS at 65°C. An additional set of stringent conditions include hybridization at 0.1X SSC, 0.1% SDS, 65°C and washes with 2X SSC, 0.1% SDS followed by 0.1X SSC, 0.1% SDS, for example.

[00120] Hybridization requires that the two nucleic acids contain complementary sequences, although depending on the stringency of the hybridization, mismatches between bases are possible. The appropriate stringency for hybridizing nucleic acids depends on the length of the nucleic acids and the degree of complementation, variables well known in the art. The greater the degree of similarity or homology between two nucleotide sequences, the greater the value of T_m for hybrids of nucleic acids having those sequences. The relative stability (corresponding to higher T_m) of nucleic acid hybridizations decreases in the following order: RNA:RNA, DNA:RNA, DNA:DNA. For hybrids of greater than 100 nucleotides in length, equations for calculating T_m have been derived (see, e.g., Sambrook et al., *supra*, 9.50-9.51). For hybridizations with shorter nucleic acids (i.e., oligonucleotides), the position of mismatches becomes more important, and the length of the oligonucleotide determines its specificity (see, e.g., Sambrook et al., *supra*, 11.7-11.8). In one embodiment, the length for a hybridizable nucleic acid is at least about 10 nucleotides. In some embodiments, a minimum length for a hybridizable nucleic acid is at least about 15 nucleotides; at least about 20 nucleotides; or the length is at least about 30 nucleotides. Furthermore, the skilled artisan

will recognize that the temperature and wash solution salt concentration may be adjusted as necessary according to factors such as length of the probe.

[00121] A "substantial portion" of an amino acid or nucleotide sequence is that portion comprising enough of the amino acid sequence of a polypeptide or the nucleotide sequence of a gene to putatively identify that polypeptide or gene, either by manual evaluation of the sequence by one skilled in the art, or by computer-automated sequence comparison and identification using algorithms such as BLAST (Altschul, et al., J. Mol. Biol. 215:403-410, 1993). In general, a sequence of ten or more contiguous amino acids or thirty or more nucleotides is necessary in order to putatively identify a polypeptide or nucleic acid sequence as homologous to a known protein or gene. Moreover, with respect to nucleotide sequences, gene specific oligonucleotide probes comprising 20-30 contiguous nucleotides may be used in sequence-dependent methods of gene identification (e.g., Southern hybridization) and isolation (e.g., *in situ* hybridization of bacterial colonies or bacteriophage plaques). In addition, short oligonucleotides of 12-15 bases may be used as amplification primers in PCR in order to obtain a particular nucleic acid fragment comprising the primers. Accordingly, a "substantial portion" of a nucleotide sequence comprises enough of the sequence to specifically identify and/or isolate a nucleic acid fragment comprising the sequence. The instant specification teaches the complete amino acid and nucleotide sequence encoding particular proteins. The skilled artisan, having the benefit of the sequences as reported herein, may now use all or a substantial portion of the disclosed sequences for purposes known to those skilled in this art. Accordingly, the instant invention comprises the complete sequences as provided herein, as well as substantial portions of those sequences as defined above.

[00122] The term "complementary" is used to describe the relationship between nucleotide bases that are capable of hybridizing to one another. For example, with respect to DNA, adenosine is complementary to thymine and cytosine is complementary to guanine.

[00123] The term "percent identity" as known in the art, is a relationship between two or more polypeptide sequences or two or more polynucleotide sequences, as determined by comparing the sequences. In the art, "identity" also means the degree of sequence relatedness between polypeptide or polynucleotide sequences, as the case may be, as determined by the match between strings of such sequences. "Identity" and "similarity" can be readily calculated by known methods including, but not limited to, those disclosed in:

Computational Molecular Biology (Lesk, A. M., Ed., Oxford University: NY, 1988); *Biocomputing: Informatics and Genome Projects* (Smith, D. W., Ed., Academic: NY, 1993); *Computer Analysis of Sequence Data, Part I* (Griffin, A. M., and Griffin, H. G., Eds. Humana: NJ, 1994); *Sequence Analysis in Molecular Biology* (von Heinje, G., Ed. Academic, 1987); and *Sequence Analysis Primer* (Gribskov, M. and Devereux, J., Eds. Stockton: NY, 1991).

[00124] Preferred methods to determine identity are designed to give the best match between the sequences tested. Methods to determine identity and similarity are codified in publicly available computer programs. Sequence alignments and percent identity calculations may be performed using the MegAlign™ program of the Lasergene® bioinformatics computing suite (DNASTAR, Inc., Madison, WI). Multiple alignment of the sequences is performed using the "Clustal method of alignment" which encompasses several varieties of the algorithm including the "Clustal V method of alignment" corresponding to the alignment method labeled Clustal V (Higgins and Sharp, CABIOS. 5:151-153, 1989; Higgins, et al., Comput. Appl. Biosci. 8:189-191, 1992) and found in the MegAlign™ program of the Lasergene® bioinformatics computing suite (DNASTAR, Inc.). For multiple alignments, the default values correspond to Gap Penalty=10 and Gap Length Penalty=10. Default parameters for pairwise alignments and calculation of percent identity of protein sequences using the Clustal method are Ktuple=1, Gap Penalty=3, Window=5 and Diagonals Saved=5. For nucleic acids these parameters are Ktuple=2, Gap Penalty=5, Window=4 and Diagonals Saved=4. After alignment of the sequences using the Clustal V program, it is possible to obtain a percent identity by viewing the sequence distances table in the same program. Additionally the "Clustal W method of alignment" is available and corresponds to the alignment method labeled Clustal W (Higgins and Sharp, CABIOS. 5:151-153, 1989; Higgins, et al., Comput. Appl. Biosci. 8:189-191, 1992) and found in the MegAlign™ v6.1 program of the Lasergene® bioinformatics computing suite (DNASTAR, Inc.). Default parameters for multiple alignment (Gap Penalty=10, Gap Length Penalty=0.2, Delay Divergen Seqs(%)=30, DNA Transition Weight=0.5, Protein Weight Matrix=Gonnet Series, DNA Weight Matrix=IUB). After alignment of the sequences using the Clustal W program, it is possible to obtain a percent identity by viewing the sequence distances table in the same program.

[00125] The term "sequence analysis software" refers to any computer algorithm or software program that is useful for the analysis of nucleotide or amino acid sequences. Sequence analysis software may be commercially available or independently developed. Sequence analysis software includes, but is not limited to: GCG suite of programs (Wisconsin Package Version 9.0, Genetics Computer Group (GCG), Madison, WI); BLASTP, BLASTN, BLASTX (Altschul, et al., J. Mol. Biol. 215:403-410, 1990); DNASTAR (DNASTAR, Inc. Madison, WI); Sequencher (Gene Codes Corporation, Ann Arbor, MI); and FASTA program incorporating the Smith-Waterman algorithm (W. R. Pearson, *Comput. Methods Genome Res.*, [Proc. Int. Symp.] (1994), Meeting Date 1992, 111-20. Editor(s): Suhai, Sandor. Plenum: New York, NY). Within the context of this application it will be understood that where sequence analysis software is used for analysis, that the results of the analysis will be based on the default values of the program referenced, unless otherwise specified. As used herein "default values" will mean any set of values or parameters that originally load with the software when first initialized.

[00126] By a nucleic acid or polynucleotide having a nucleotide sequence at least, for example, 95% identical to a reference nucleotide sequence of the present invention, it is intended that the nucleotide sequence of the polynucleotide is identical to the reference sequence except that the polynucleotide sequence may include up to five point mutations per each 100 nucleotides of the reference nucleotide sequence. In other words, to obtain a polynucleotide having a nucleotide sequence at least 95% identical to a reference nucleotide sequence, up to 5% of the nucleotides in the reference sequence may be deleted or substituted with another nucleotide, or a number of nucleotides up to 5% of the total nucleotides in the reference sequence may be inserted into the reference sequence.

[00127] As a practical matter, whether any particular nucleic acid molecule or polypeptide is at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identical to a nucleotide sequence or polypeptide sequence of the present invention can be determined conventionally using known computer programs. A preferred method for determining the best overall match between a query sequence (e.g., a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag, et al., (Comp. Appl. Biosci. 6:237-245, 1990). In a sequence alignment, the query and subject sequences are both DNA sequences. An RNA sequence can be compared by converting U's to T's. The result of the global

sequence alignment is in percent identity. Preferred parameters used in a FASTDB alignment of DNA sequences to calculate percent identity are: Matrix=Unitary, k-tuple=4, Mismatch Penalty=1, Joining Penalty=30, Randomization Group Length=0, Cutoff Score=1, Gap Penalty=5, Gap Size Penalty=0.05, Window Size=500 or the length of the subject nucleotide sequences, whichever is shorter.

[00128] If the subject sequence is shorter than the query sequence because of 5' or 3' deletions, not because of internal deletions, a manual correction must be made to the results. This is because the FASTDB program does not account for 5' and 3' truncations of the subject sequence when calculating percent identity. For subject sequences truncated at the 5' or 3' ends, relative to the query sequence, the percent identity is corrected by calculating the number of bases of the query sequence that are 5' and 3' of the subject sequence, which are not matched/aligned, as a percent of the total bases of the query sequence. Whether a nucleotide is matched/aligned is determined by results of the FASTDB sequence alignment. This percentage is then subtracted from the percent identity, calculated by the above FASTDB program using the specified parameters, to arrive at a final percent identity score. This corrected score is what is used for the purposes of the present invention. Only bases outside the 5' and 3' bases of the subject sequence, as displayed by the FASTDB alignment, which are not matched/aligned with the query sequence, are calculated for the purposes of manually adjusting the percent identity score.

[00129] For example, a 90 base subject sequence is aligned to a 100 base query sequence to determine percent identity. The deletions occur at the 5' end of the subject sequence and therefore, the FASTDB alignment does not show a matched/alignment of the first 10 bases at 5' end. The 10 unpaired bases represent 10% of the sequence (number of bases at the 5' and 3' ends not matched/total number of bases in the query sequence) so 10% is subtracted from the percent identity score calculated by the FASTDB program. If the remaining 90 bases were perfectly matched the final percent identity would be 90%. In another example, a 90 base subject sequence is compared with a 100 base query sequence. This time the deletions are internal deletions so that there are no bases on the 5' or 3' of the subject sequence which are not matched/aligned with the query. In this case, the percent identity calculated by FASTDB is not manually corrected. Once again, only bases 5' and 3' of the subject sequence which are not matched/aligned with the query sequence are manually corrected for. No other manual corrections are to be made for the purposes of the present invention.

[00130] Standard recombinant DNA and molecular cloning techniques used here are well known in the art and are described by Sambrook, J., Fritsch, E. F. and Maniatis, T., *Molecular Cloning: A Laboratory Manual*, Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1989); Silhavy, T. J., Bennis, M. L. and Enquist, L. W., *Experiments with Gene Fusions*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1984); and Ausubel, F. M. *et al.*, *Current Protocols in Molecular Biology*, published by Greene Publishing Assoc. and Wiley-Interscience (1987). Additional methods used include in *Methods in Enzymology*, Volume 194, *Guide to Yeast Genetics and Molecular and Cell Biology* (Part A, 2004, Christine Guthrie and Gerald R. Fink (Eds.), Elsevier Academic Press, San Diego, CA).

[00131] Methods for increasing or for reducing gene expression of the target genes above are well known to one skilled in the art. Methods for gene expression in yeasts are known in the art as described, for example, in *Methods in Enzymology*, Volume 194, *Guide to Yeast Genetics and Molecular and Cell Biology* (Part A, 2004, Christine Guthrie and Gerald R. Fink (Eds.), Elsevier Academic Press, San Diego, Calif.). For example, methods for increasing expression include increasing the number of genes that are integrated in the genome or on plasmids that express the target protein, and using a promoter that is more highly expressed than the natural promoter. Promoters that may be operably linked in a constructed chimeric gene for expression include, for example, constitutive promoters FBA1, TDH3, ADH1, and GPM1, and the inducible promoters GAL1, GAL10, and CUP1. Suitable transcriptional terminators that may be used in a chimeric gene construct for expression include, but are not limited to FBA1t, TDH3t, GPM1t, ERG10t, GAL1t, CYC1t, and ADH1t.

[00132] Suitable promoters, transcriptional terminators, and coding regions may be cloned into *E. coli*-yeast shuttle vectors, and transformed into yeast cells. These vectors allow for propagation in both *E. coli* and yeast strains. Typically, the vector contains a selectable marker and sequences allowing autonomous replication or chromosomal integration in the desired host. Plasmids used in yeast are, for example, shuttle vectors pRS423, pRS424, pRS425, and pRS426 (American Type Culture Collection, Rockville, Md.), which contain an *E. coli* replication origin (e.g., pMB1), a yeast 2 μ origin of replication, and a marker for nutritional selection. The selection markers for these four vectors are HIS3 (vector pRS423), TRP1 (vector pRS424), LEU2 (vector pRS425), and URA3 (vector pRS426). Construction

of expression vectors may be performed by either standard molecular cloning techniques in *E. coli* or by the gap repair recombination method in yeast.

[00133] Methods for reducing expression include using genetic modification of the encoding genes. Many methods for genetic modification of target genes to reduce or eliminate expression are known to one skilled in the art and may be used to create the present yeast production host cells. Modifications that may be used include, but are not limited to, deletion of the entire gene or a portion of the gene encoding the protein, inserting a DNA fragment into the encoding gene (in either the promoter or coding region) so that the protein is not expressed or expressed at lower levels, introducing a mutation into the coding region which adds a stop codon or frame shift such that a functional protein is not expressed, and introducing one or more mutations into the coding region to alter amino acids so that a non-functional or a less active protein is expressed. In addition, expression of a target gene may be blocked by expression of an antisense RNA or an interfering RNA, and constructs may be introduced that result in cosuppression. In addition, the synthesis or stability of the transcript may be lessened by mutation. Similarly, the efficiency by which a protein is translated from mRNA may be modulated by mutation. All of these methods may be readily practiced by one skilled in the art making use of the known or identified sequences encoding target proteins.

[00134] DNA sequences surrounding a target coding sequence are also useful in some modification procedures. In particular, DNA sequences surrounding, for example, a target gene coding sequence are useful for modification methods using homologous recombination. For example, in this method target gene flanking sequences are placed bounding a selectable marker gene to mediate homologous recombination whereby the marker gene replaces the target gene. Also, partial target gene sequences and target gene flanking sequences bounding a selectable marker gene may be used to mediate homologous recombination whereby the marker gene replaces a portion of the target gene. In addition, the selectable marker may be bounded by site-specific recombination sites, so that following expression of the corresponding site-specific recombinase, the resistance gene is excised from the target gene without reactivating the latter. The site-specific recombination leaves behind a recombination site which disrupts expression of the target protein. The homologous recombination vector may be constructed to also leave a deletion in the target gene following excision of the selectable marker, as is well known to one skilled in the art.

[00135] Deletions may be made using mitotic recombination as described in Wach, et al. (Yeast 10:1793-1808, 1994). This method involves preparing a DNA fragment that contains a selectable marker between genomic regions that may be as short as 20 bp, and which bound a target DNA sequence. This DNA fragment can be prepared by PCR amplification of the selectable marker gene using as primers oligonucleotides that hybridize to the ends of the marker gene and that include the genomic regions that can recombine with the yeast genome. The linear DNA fragment can be efficiently transformed into yeast and recombined into the genome resulting in gene replacement including with deletion of the target DNA sequence (Methods in Enzymology, v 194, pp 281-301, 1991).

[00136] Moreover, promoter replacement methods may be used to exchange the endogenous transcriptional control elements allowing another means to modulate expression (see, e.g., Mnaimneh, et al., Cell 118:31-44, 2004).

[00137] In addition, target gene encoded activity may be disrupted using random mutagenesis, which is followed by screening to identify strains with reduced activity. Using this type of method, the DNA sequence of the target gene encoding region, or any other region of the genome affecting activity, need not be known. Methods for creating genetic mutations are common and well known in the art and may be applied to the exercise of creating mutants. Commonly used random genetic modification methods (reviewed in *Methods in Yeast Genetics*, 2005, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.) include spontaneous mutagenesis, mutagenesis caused by mutator genes, chemical mutagenesis, irradiation with UV or X-rays, or transposon mutagenesis.

[00138] Chemical mutagenesis of yeast commonly involves treatment of yeast cells with one of the following DNA mutagens: ethyl methanesulfonate (EMS), nitrous acid, diethyl sulfate, or N-methyl-N'-nitro-N-nitroso-guanidine (MNNG). These methods of mutagenesis have been reviewed in Spencer, et al. (Mutagenesis in Yeast, *Yeast Protocols: Methods in Cell and Molecular Biology*. Humana Press, Totowa, N.J., 1996). Chemical mutagenesis with EMS may be performed as described in *Methods in Yeast Genetics* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 2005). Irradiation with ultraviolet (UV) light or X-rays can also be used to produce random mutagenesis in yeast cells. The primary effect of mutagenesis by UV irradiation is the formation of pyrimidine dimers which disrupt the fidelity of DNA replication. Protocols for UV-mutagenesis of yeast can be found in Spencer, et al. (Mutagenesis in Yeast, *Yeast Protocols: Methods in Cell and Molecular Biology*.

Humana Press, Totowa, N.J., 1996). Introduction of a mutator phenotype can also be used to generate random chromosomal mutations in yeast. Common mutator phenotypes can be obtained through disruption of one or more of the following genes: PMS1, MAG1, RAD18 or RAD51. Restoration of the non-mutator phenotype can be easily obtained by insertion of the wild type allele.

[00139] Many methods for genetic modification of target genes to increase, reduce, or eliminate expression are known to one of ordinary skill in the art and may be used to create a recombinant host cell disclosed herein. Further, modifications of a target gene in a recombinant host cell disclosed herein may be confirmed using methods known in the art. For example, disruption of a target may be confirmed with PCR screening using primers internal and external to the gene or by Southern blot using a probe designed to the gene sequence.

Biosynthetic Pathways

[00140] Biosynthetic pathways for the production of isobutanol that may be used include those described in U.S. Patent No. 7,851,188, the entire contents of which are herein incorporated by reference. In one embodiment, the isobutanol biosynthetic pathway comprises the following substrate to product conversions:

- a) pyruvate to acetolactate, which may be catalyzed, for example, by acetolactate synthase;
- b) acetolactate to 2,3-dihydroxyisovalerate, which may be catalyzed, for example, by acetohydroxy acid reductoisomerase;
- c) 2,3-dihydroxyisovalerate to α -ketoisovalerate, which may be catalyzed, for example, by acetohydroxy acid dehydratase;
- d) α -ketoisovalerate to isobutyraldehyde, which may be catalyzed, for example, by a branched-chain α -keto acid decarboxylase; and
- e) isobutyraldehyde to isobutanol, which may be catalyzed, for example, by a branched-chain alcohol dehydrogenase.

[00141] In another embodiment, the isobutanol biosynthetic pathway comprises the following substrate to product conversions:

- a) pyruvate to acetolactate, which may be catalyzed, for example, by acetolactate synthase;

- 42 -

- b) acetolactate to 2,3-dihydroxyisovalerate, which may be catalyzed, for example, by ketol-acid reductoisomerase;
- c) 2,3-dihydroxyisovalerate to α -ketoisovalerate, which may be catalyzed, for example, by dihydroxyacid dehydratase;
- d) α -ketoisovalerate to valine, which may be catalyzed, for example, by transaminase or valine dehydrogenase;
- e) valine to isobutylamine, which may be catalyzed, for example, by valine decarboxylase;
- f) isobutylamine to isobutyraldehyde, which may be catalyzed by, for example, omega transaminase; and
- g) isobutyraldehyde to isobutanol, which may be catalyzed, for example, by a branched-chain alcohol dehydrogenase.

[00142] In another embodiment, the isobutanol biosynthetic pathway comprises the following substrate to product conversions:

- a) pyruvate to acetolactate, which may be catalyzed, for example, by acetolactate synthase;
- b) acetolactate to 2,3-dihydroxyisovalerate, which may be catalyzed, for example, by acetohydroxy acid reductoisomerase;
- c) 2,3-dihydroxyisovalerate to α -ketoisovalerate, which may be catalyzed, for example, by acetohydroxy acid dehydratase;
- d) α -ketoisovalerate to isobutyryl-CoA, which may be catalyzed, for example, by branched-chain keto acid dehydrogenase;
- e) isobutyryl-CoA to isobutyraldehyde, which may be catalyzed, for example, by acetylating aldehyde dehydrogenase; and
- f) isobutyraldehyde to isobutanol, which may be catalyzed, for example, by a branched-chain alcohol dehydrogenase.

[00143] Biosynthetic pathways for the production of 1-butanol that may be used include those described in U.S. Patent Application Publication No. 2008/0182308, the entire contents of which are herein incorporated by reference. In one embodiment, the 1-butanol biosynthetic pathway comprises the following substrate to product conversions:

- a) acetyl-CoA to acetoacetyl-CoA, which may be catalyzed, for example, by acetyl-CoA acetyltransferase;

- 43 -

- b) acetoacetyl-CoA to 3-hydroxybutyryl-CoA, which may be catalyzed, for example, by 3-hydroxybutyryl-CoA dehydrogenase;
- c) 3-hydroxybutyryl-CoA to crotonyl-CoA, which may be catalyzed, for example, by crotonase;
- d) crotonyl-CoA to butyryl-CoA, which may be catalyzed, for example, by butyryl-CoA dehydrogenase;
- e) butyryl-CoA to butyraldehyde, which may be catalyzed, for example, by butyraldehyde dehydrogenase; and
- f) butyraldehyde to 1-butanol, which may be catalyzed, for example, by butanol dehydrogenase.

[00144] Biosynthetic pathways for the production of 2-butanol that may be used include those described in U.S. Patent Application Publication No. 2007/0259410 and U.S. Patent Application Publication No. 2009/0155870, the entire contents of each are herein incorporated by reference. In one embodiment, the 2-butanol biosynthetic pathway comprises the following substrate to product conversions:

- a) pyruvate to alpha-acetolactate, which may be catalyzed, for example, by acetolactate synthase;
- b) alpha-acetolactate to acetoin, which may be catalyzed, for example, by acetolactate decarboxylase;
- c) acetoin to 3-amino-2-butanol, which may be catalyzed, for example, acetoin aminase;
- d) 3-amino-2-butanol to 3-amino-2-butanol phosphate, which may be catalyzed, for example, by aminobutanol kinase;
- e) 3-amino-2-butanol phosphate to 2-butanone, which may be catalyzed, for example, by aminobutanol phosphate phosphorylase; and
- f) 2-butanone to 2-butanol, which may be catalyzed, for example, by butanol dehydrogenase.

[00145] In another embodiment, the 2-butanol biosynthetic pathway comprises the following substrate to product conversions:

- a) pyruvate to alpha-acetolactate, which may be catalyzed, for example, by acetolactate synthase;

- 44 -

- b) alpha-acetolactate to acetoin, which may be catalyzed, for example, by acetolactate decarboxylase;
- c) acetoin to 2,3-butanediol, which may be catalyzed, for example, by butanediol dehydrogenase;
- d) 2,3-butanediol to 2-butanone, which may be catalyzed, for example, by dialdehydratase; and
- e) 2-butanone to 2-butanol, which may be catalyzed, for example, by butanol dehydrogenase.

[00146] Biosynthetic pathways for the production of 2-butanone that may be used include those described in U.S. Patent Application Publication No. 2007/0259410 and U.S. Patent Application Publication No. 2009/0155870, the entire contents of each are herein incorporated by reference. In one embodiment, the 2-butanone biosynthetic pathway comprises the following substrate to product conversions:

- a) pyruvate to alpha-acetolactate, which may be catalyzed, for example, by acetolactate synthase;
- b) alpha-acetolactate to acetoin, which may be catalyzed, for example, by acetolactate decarboxylase;
- c) acetoin to 3-amino-2-butanol, which may be catalyzed, for example, acetoin aminase;
- d) 3-amino-2-butanol to 3-amino-2-butanol phosphate, which may be catalyzed, for example, by aminobutanol kinase; and
- e) 3-amino-2-butanol phosphate to 2-butanone, which may be catalyzed, for example, by aminobutanol phosphate phosphorylase.

[00147] In another embodiment, the 2-butanone biosynthetic pathway comprises the following substrate to product conversions:

- a) pyruvate to alpha-acetolactate, which may be catalyzed, for example, by acetolactate synthase;
- b) alpha-acetolactate to acetoin which may be catalyzed, for example, by acetolactate decarboxylase;
- c) acetoin to 2,3-butanediol, which may be catalyzed, for example, by butanediol dehydrogenase; and

- d) 2,3-butanediol to 2-butanone, which may be catalyzed, for example, by diol dehydratase.

[00148] In one embodiment, the invention produces butanol from plant-derived carbon sources, avoiding the negative environmental impact associated with standard petrochemical processes for butanol production. In one embodiment, the invention provides a method for the production of butanol using recombinant industrial host cells comprising a butanol pathway.

[00149] In some embodiments, the isobutanol biosynthetic pathway comprises at least one polynucleotide, at least two polynucleotides, at least three polynucleotides, at least four polynucleotides, or more that is/are heterologous to the host cell. In some embodiments, each substrate to product conversion of an isobutanol biosynthetic pathway in a recombinant host cell is catalyzed by a heterologous polypeptide. In some embodiments, the polypeptide catalyzing the substrate to product conversions of acetolactate to 2,3-dihydroxyisovalerate and/or the polypeptide catalyzing the substrate to product conversion of isobutyraldehyde to isobutanol are capable of utilizing NADH as a cofactor.

[00150] The terms "acetohydroxyacid synthase," "acetolactate synthase," and "acetolactate synthetase" (abbreviated "ALS") may be used interchangeably herein to refer to a polypeptide having enzymatic activity that catalyzes the conversion of pyruvate to acetolactate and CO₂. Example acetolactate synthases are known by the EC number 2.2.1.6 (Enzyme Nomenclature 1992, Academic Press, San Diego). These unmodified enzymes are available from a number of sources, including, but not limited to, *Bacillus subtilis* (GenBank Nos: CAB15618 (SEQ ID NO: 1), Z99122 (SEQ ID NO: 2), NCBI (National Center for Biotechnology Information) amino acid sequence, NCBI nucleotide sequence, respectively), *Klebsiella pneumoniae* (GenBank Nos: AAA25079 (SEQ ID NO: 3), M73842 (SEQ ID NO: 4)), and *Lactococcus lactis* (GenBank Nos: AAA25161 (SEQ ID NO: 5), L16975 (SEQ ID NO: 6)).

[00151] The terms "ketol-acid reductoisomerase" ("KARI"), "acetohydroxy acid isomeroreductase," and "acetohydroxy acid reductoisomerase" will be used interchangeably and refer to a polypeptide having enzymatic activity capable of catalyzing the reaction of (S)-acetolactate to 2,3-dihydroxyisovalerate. Example KARI enzymes may be classified as EC number EC 1.1.1.86 (Enzyme Nomenclature 1992, Academic Press, San Diego), and are available from a vast array of microorganisms including, but not limited to, *Escherichia coli*

(GenBank Nos: NP_418222 (SEQ ID NO: 7), NC_000913 (SEQ ID NO: 8)), *Saccharomyces cerevisiae* (GenBank Nos: NP_013459 (SEQ ID NO: 9), NC_001144 (SEQ ID NO: 10)), *Methanococcus maripaludis* (GenBank Nos: CAF30210 (SEQ ID NO: 11), BX957220 (SEQ ID NO: 12)), and *Bacillus subtilis* (GenBank Nos: CAB14789 (SEQ ID NO: 13), Z99118 (SEQ ID NO: 14)). KARIs include *Anaerostipes caccae* KARI variants "K9G9" and "K9D3" (SEQ ID NOs: 15 and 16, respectively). Ketol-acid reductoisomerase (KARI) enzymes are described in U.S. Patent Application Publication Nos. 2008/0261230, 2009/0163376, and 2010/0197519, and PCT Application Publication No. WO/2011/041415, the entire contents of each are herein incorporated by reference. Examples of KARIs disclosed therein are those from *Lactococcus lactis*, *Vibrio cholera*, *Pseudomonas aeruginosa* PAO1, and *Pseudomonas fluorescens* PF5 mutants. In some embodiments, the KARI utilizes NADH. In some embodiments, the KARI utilizes NADPH.

[00152] The terms "acetohydroxy acid dehydratase" and "dihydroxyacid dehydratase" ("DHAD") refers to a polypeptide having enzymatic activity that catalyzes the conversion of 2,3-dihydroxyisovalerate to α -ketoisovalerate. Example acetohydroxy acid dehydratases are known by the EC number 4.2.1.9. Such enzymes are available from a vast array of microorganisms including, but not limited to, *E. coli* (GenBank Nos: YP_026248 (SEQ ID NO: 17), NC000913 (SEQ ID NO: 18)), *Saccharomyces cerevisiae* (GenBank Nos: NP_012550 (SEQ ID NO: 19), NC_001142 (SEQ ID NO: 20)), *M. maripaludis* (GenBank Nos: CAF29874 (SEQ ID NO: 21), BX957219 (SEQ ID NO: 22)), *B. subtilis* (GenBank Nos: CAB14105 (SEQ ID NO: 23), Z99115 (SEQ ID NO: 24)), *L. lactis*, and *N. crassa*. U.S. Patent Application Publication No. 2010/0081154, and U.S. Patent No. 7,851,188, the entire contents of each are herein incorporated by reference, describe dihydroxyacid dehydratases (DHADs), including a DHAD from *Streptococcus mutans*.

[00153] The terms "branched-chain α -keto acid decarboxylase," " α -ketoacid decarboxylase," " α -ketoisovalerate decarboxylase," or "2-ketoisovalerate decarboxylase" ("KIVD") refers to a polypeptide having enzymatic activity that catalyzes the conversion of α -ketoisovalerate to isobutyraldehyde and CO₂. Example branched-chain α -keto acid decarboxylases are known by the EC number 4.1.1.72 and are available from a number of sources including, but not limited to, *Lactococcus lactis* (GenBank Nos: AAS49166 (SEQ ID NO: 25), AY548760 (SEQ ID NO: 26); CAG34226 (SEQ ID NO: 27), AJ746364 (SEQ ID NO: 28), *Salmonella typhimurium* (GenBank Nos: NP_461346 (SEQ ID NO: 29),

NC_003197 (SEQ ID NO: 30)), *Clostridium acetobutylicum* (GenBank Nos: NP_149189 (SEQ ID NO: 31), NC_001988 (SEQ ID NO: 32)), *M. caseolyticus* (SEQ ID NO: 33), and *L. grayi* (SEQ ID NO: 34).

[00154] The term "branched-chain alcohol dehydrogenase" ("ADH") refers to a polypeptide having enzymatic activity that catalyzes the conversion of isobutyraldehyde to isobutanol. Example branched-chain alcohol dehydrogenases are known by the EC number 1.1.1.265, but may also be classified under other alcohol dehydrogenases (specifically, EC 1.1.1.1 or 1.1.1.2). Alcohol dehydrogenases may be NADPH-dependent or NADH-dependent. Such enzymes are available from a number of sources including, but not limited to, *S. cerevisiae* (GenBank Nos: NP_010656 (SEQ ID NO: 35), NC_001136 (SEQ ID NO: 36), NP_014051 (SEQ ID NO: 37), NC_001145 (SEQ ID NO: 38)), *E. coli* (GenBank Nos: NP_417484 (SEQ ID NO: 39), NC_000913 (SEQ ID NO: 40)), *C. acetobutylicum* (GenBank Nos: NP_349892 (SEQ ID NO: 41), NC_003030 (SEQ ID NO: 42); NP_349891 (SEQ ID NO: 43), NC_003030 (SEQ ID NO: 44)). U.S. Patent Application Publication No. 2009/0269823 describes SadB, an alcohol dehydrogenase (ADH) from *Achromobacter xylosoxidans*. Alcohol dehydrogenases also include horse liver ADH and *Beijerinckia indica* ADH (as described by U.S. Patent Application Publication No. 2011/0269199, the entire contents of which are herein incorporated by reference).

[00155] The term "butanol dehydrogenase" refers to a polypeptide having enzymatic activity that catalyzes the conversion of isobutyraldehyde to isobutanol or the conversion of 2-butanone and 2-butanol. Butanol dehydrogenases are a subset of a broad family of alcohol dehydrogenases. Butanol dehydrogenase may be NAD-dependent or NADP-dependent. The NAD-dependent enzymes are known as EC 1.1.1.1 and are available, for example, from *Rhodococcus ruber* (GenBank Nos: CAD36475, AJ491307). The NADP-dependent enzymes are known as EC 1.1.1.2 and are available, for example, from *Pyrococcus furiosus* (GenBank Nos: AAC25556, AF013169). Additionally, a butanol dehydrogenase is available from *Escherichia coli* (GenBank Nos: NP_417484, NC_000913) and a cyclohexanol dehydrogenase is available from *Acinetobacter sp.* (GenBank Nos: AAG10026, AF282240). The term "butanol dehydrogenase" also refers to a polypeptide having enzymatic activity that catalyzes the conversion of butyraldehyde to 1-butanol, using either NADH or NADPH as cofactor. Butanol dehydrogenases are available from, for example, *C. acetobutylicum* (GenBank Nos: NP_149325, NC_001988; this enzyme possesses both aldehyde and alcohol

dehydrogenase activity); NP_349891, NC_003030; and NP_349892, NC_003030) and *E. coli* (GenBank Nos: NP_417-484, NC_000913).

[00156] The term "branched-chain keto acid dehydrogenase" refers to a polypeptide having enzymatic activity that catalyzes the conversion of α -ketoisovalerate to isobutyryl-CoA (isobutyryl-coenzyme A), typically using NAD⁺ (nicotinamide adenine dinucleotide) as an electron acceptor. Example branched-chain keto acid dehydrogenases are known by the EC number 1.2.4.4. Such branched-chain keto acid dehydrogenases are comprised of four subunits and sequences from all subunits are available from a vast array of microorganisms including, but not limited to, *B. subtilis* (GenBank Nos: CAB14336 (SEQ ID NO: 45), Z99116 (SEQ ID NO: 46); CAB14335 (SEQ ID NO: 47), Z99116 (SEQ ID NO: 48); CAB14334 (SEQ ID NO: 49), Z99116 (SEQ ID NO: 50); and CAB14337 (SEQ ID NO: 51), Z99116 (SEQ ID NO: 52)) and *Pseudomonas putida* (GenBank Nos: AAA65614 (SEQ ID NO: 53), M57613 (SEQ ID NO: 54); AAA65615 (SEQ ID NO: 55), M57613 (SEQ ID NO: 56); AAA65617 (SEQ ID NO: 57), M57613 (SEQ ID NO: 58); and AAA65618 (SEQ ID NO: 59), M57613 (SEQ ID NO: 60)).

[00157] The term "acylating aldehyde dehydrogenase" refers to a polypeptide having enzymatic activity that catalyzes the conversion of isobutyryl-CoA to isobutyraldehyde, typically using either NADH or NADPH as an electron donor. Example acylating aldehyde dehydrogenases are known by the EC numbers 1.2.1.10 and 1.2.1.57. Such enzymes are available from multiple sources including, but not limited to, *Clostridium beijerinckii* (GenBank Nos: AAD31841 (SEQ ID NO: 61), AF157306 (SEQ ID NO: 62)), *C. acetobutylicum* (GenBank Nos: NP_149325 (SEQ ID NO: 63), NC_001988 (SEQ ID NO: 64); NP_149199 (SEQ ID NO: 65), NC_001988 (SEQ ID NO: 66)), *P. putida* (GenBank Nos: AAA89106 (SEQ ID NO: 67), U13232 (SEQ ID NO: 68)), and *Thermus thermophilus* (GenBank Nos: YP_145486 (SEQ ID NO: 69), NC_006461 (SEQ ID NO: 70)).

[00158] The term "transaminase" refers to a polypeptide having enzymatic activity that catalyzes the conversion of α -ketoisovalerate to L-valine, using either alanine or glutamate as an amine donor. Example transaminases are known by the EC numbers 2.6.1.42 and 2.6.1.66. Such enzymes are available from a number of sources. Examples of sources for alanine-dependent enzymes include, but are not limited to, *E. coli* (GenBank Nos: YP_026231 (SEQ ID NO: 71), NC_000913 (SEQ ID NO: 72)) and *Bacillus licheniformis* (GenBank Nos: YP_093743 (SEQ ID NO: 73), NC_006322 (SEQ ID NO: 74)). Examples of

sources for glutamate-dependent enzymes include, but are not limited to, *E. coli* (GenBank Nos: YP_026247 (SEQ ID NO: 75), NC_000913 (SEQ ID NO: 76)), *Saccharomyces cerevisiae* (GenBank Nos: NP_012682 (SEQ ID NO: 77), NC_001142 (SEQ ID NO: 78)) and *Methanobacterium thermoautotrophicum* (GenBank Nos: NP_276546 (SEQ ID NO: 79), NC_000916 (SEQ ID NO: 80)).

[00159] The term "valine dehydrogenase" refers to a polypeptide having enzymatic activity that catalyzes the conversion of α -ketoisovalerate to L-valine, typically using NADPH as an electron donor and ammonia as an amine donor. Example valine dehydrogenases are known by the EC numbers 1.4.1.8 and 1.4.1.9 and such enzymes are available from a number of sources including, but not limited to, *Streptomyces coelicolor* (GenBank Nos: NP_628270 (SEQ ID NO: 81), NC_003888 (SEQ ID NO: 82)) and *B. subtilis* (GenBank Nos: CAB14339 (SEQ ID NO: 83), Z99116 (SEQ ID NO: 84)).

[00160] The term "valine decarboxylase" refers to a polypeptide having enzymatic activity that catalyzes the conversion of L-valine to isobutylamine and CO₂. Example valine decarboxylases are known by the EC number 4.1.1.14. Such enzymes are found in *Streptomyces*, such as for example, *Streptomyces viridifaciens* (GenBank Nos: AAN10242 (SEQ ID NO: 85), AY116644 (SEQ ID NO: 86)).

[00161] The term "omega transaminase" refers to a polypeptide having enzymatic activity that catalyzes the conversion of isobutylamine to isobutyraldehyde using a suitable amino acid as an amine donor. Example omega transaminases are known by the EC number 2.6.1.18 and are available from a number of sources including, but not limited to, *Alcaligenes denitrificans* (AAP92672 (SEQ ID NO: 87), AY330220 (SEQ ID NO: 88)), *Ralstonia eutropha* (GenBank Nos: YP_294474 (SEQ ID NO: 89), NC_007347 (SEQ ID NO: 90)), *Shewanella oneidensis* (GenBank Nos: NP_719046 (SEQ ID NO: 91), NC_004347 (SEQ ID NO: 92)), and *P. putida* (GenBank Nos: AAN66223 (SEQ ID NO: 93), AE016776 (SEQ ID NO: 94)).

[00162] The term "acetyl-CoA acetyltransferase" refers to a polypeptide having enzymatic activity that catalyzes the conversion of two molecules of acetyl-CoA to acetoacetyl-CoA and coenzyme A (CoA). Example acetyl-CoA acetyltransferases are acetyl-CoA acetyltransferases with substrate preferences (reaction in the forward direction) for a short chain acyl-CoA and acetyl-CoA and are classified as E.C. 2.3.1.9 [Enzyme Nomenclature 1992, Academic Press, San Diego]; although, enzymes with a broader substrate range (E.C.

2.3.1.16) will be functional as well. Acetyl-CoA acetyltransferases are available from a number of sources, for example, *Escherichia coli* (GenBank Nos: NP_416728, NC_000913; NCBI amino acid sequence, NCBI nucleotide sequence), *Clostridium acetobutylicum* (GenBank Nos: NP_349476.1, NC_003030; NP_149242, NC_001988, *Bacillus subtilis* (GenBank Nos: NP_390297, NC_000964), and *Saccharomyces cerevisiae* (GenBank Nos: NP_015297, NC_001148).

[00163] The term "3-hydroxybutyryl-CoA dehydrogenase" refers to a polypeptide having enzymatic activity that catalyzes the conversion of acetoacetyl-CoA to 3-hydroxybutyryl-CoA. Example hydroxybutyryl-CoA dehydrogenases may be NADH-dependent, with a substrate preference for (S)-3-hydroxybutyryl-CoA or (R)-3-hydroxybutyryl-CoA. Examples may be classified as E.C. 1.1.1.35 and E.C. 1.1.1.30, respectively. Additionally, 3-hydroxybutyryl-CoA dehydrogenases may be NADPH-dependent, with a substrate preference for (S)-3-hydroxybutyryl-CoA or (R)-3-hydroxybutyryl-CoA and are classified as E.C. 1.1.1.157 and E.C. 1.1.1.36, respectively. 3-Hydroxybutyryl-CoA dehydrogenases are available from a number of sources, for example, *C. acetobutylicum* (GenBank Nos: NP_349314, NC_003030), *B. subtilis* (GenBank Nos: AAB09614, U29084), *Ralstonia eutropha* (GenBank Nos: YP_294481, NC_007347), and *Alcaligenes eutrophus* (GenBank Nos: AAA21973, J04987).

[00164] The term "crotonase" refers to a polypeptide having enzymatic activity that catalyzes the conversion of 3-hydroxybutyryl-CoA to crotonyl-CoA and H₂O. Example crotonases may have a substrate preference for (S)-3-hydroxybutyryl-CoA or (R)-3-hydroxybutyryl-CoA and may be classified as E.C. 4.2.1.17 and E.C. 4.2.1.55, respectively. Crotonases are available from a number of sources, for example, *E. coli* (GenBank Nos: NP_415911, NC_000913), *C. acetobutylicum* (GenBank Nos: NP_349318, NC_003030), *B. subtilis* (GenBank Nos: CAB13705, Z99113), and *Aeromonas caviae* (GenBank Nos: BAA21816, D88825).

[00165] The term "butyryl-CoA dehydrogenase" refers to a polypeptide having enzymatic activity that catalyzes the conversion of crotonyl-CoA to butyryl-CoA. Example butyryl-CoA dehydrogenases may be NADH-dependent, NADPH-dependent, or flavin-dependent and may be classified as E.C. 1.3.1.44, E.C. 1.3.1.38, and E.C. 1.3.99.2, respectively. Butyryl-CoA dehydrogenases are available from a number of sources, for example, *C. acetobutylicum* (GenBank Nos: NP_347102, NC_003030), *Euglena gracilis* (GenBank Nos:

Q5EU90), AY741582), *Streptomyces collinus* (GenBank Nos: AAA92890, U37135), and *Streptomyces coelicolor* (GenBank Nos: CAA22721, AL939127).

[00166] The term "butyraldehyde dehydrogenase" refers to a polypeptide having enzymatic activity that catalyzes the conversion of butyryl-CoA to butyraldehyde, using NADH or NADPH as cofactor. Butyraldehyde dehydrogenases with a preference for NADH are known as E.C. 1.2.1.57 and are available from, for example, *Clostridium beijerinckii* (GenBank Nos: AAD31841, AF157306) and *C. acetobutylicum* (GenBank Nos: NP.sub.--149325, NC.sub.--001988).

[00167] The term "isobutyryl-CoA mutase" refers to a polypeptide having enzymatic activity that catalyzes the conversion of butyryl-CoA to isobutyryl-CoA. This enzyme may use coenzyme B₁₂ as cofactor. Example isobutyryl-CoA mutases are known by the EC number 5.4.99.13. These enzymes are found in a number of *Streptomyces* including, but not limited to, *Streptomyces cinnamonensis* (GenBank Nos: AAC08713 (SEQ ID NO: 95), U67612 (SEQ ID NO: 96); CAB59633 (SEQ ID NO: 97), AJ246005 (SEQ ID NO: 98)), *S. coelicolor* (GenBank Nos: CAB70645 (SEQ ID NO: 99), AL939123 (SEQ ID NO: 100); CAB92663 (SEQ ID NO: 101), AL939121 (SEQ ID NO: 102)), and *Streptomyces avermitilis* (GenBank Nos: NP_824008 (SEQ ID NO: 103), NC_003155 (SEQ ID NO: 104); NP_824637 (SEQ ID NO: 105), NC_003155 (SEQ ID NO: 106)).

[00168] The term "acetolactate decarboxylase" refers to a polypeptide having enzymatic activity that catalyzes the conversion of alpha-acetolactate to acetoin. Example acetolactate decarboxylases are known as EC 4.1.1.5 and are available, for example, from *Bacillus subtilis* (GenBank Nos: AAA22223, L04470), *Klebsiella terrigena* (GenBank Nos: AAA25054, L04507) and *Klebsiella pneumoniae* (GenBank Nos: AAU43774, AY722056).

[00169] The terms "acetoin aminase" or "acetoin transaminase" refers to a polypeptide having enzymatic activity that catalyzes the conversion of acetoin to 3-amino-2-butanol. Acetoin aminase may utilize the cofactor pyridoxal 5'-phosphate, NADH, or NADPH. The resulting product may have (R)- or (S)-stereochemistry at the 3-position. The pyridoxal phosphate-dependent enzyme may use an amino acid such as alanine or glutamate as the amino donor. The NADH-dependent and NADPH-dependent enzymes may use ammonia as a second substrate. A suitable example of an NADH-dependent acetoin aminase, also known as amino alcohol dehydrogenase, is described by Ito, et al. (U.S. Patent No. 6,432,688). An example of a pyridoxal-dependent acetoin aminase is the amine:pyruvate aminotransferase

(also called amine:pyruvate transaminase) described by Shin and Kim (J. Org. Chem. 67:2848-2853, 2002).

[00170] The term "acetoin kinase" refers to a polypeptide having enzymatic activity that catalyzes the conversion of acetoin to phosphoacetoin. Acetoin kinase may utilize ATP (adenosine triphosphate) or phosphoenolpyruvate as the phosphate donor in the reaction. Enzymes that catalyze the analogous reaction on the similar substrate dihydroxyacetone, for example, include enzymes known as EC 2.7.1.29 (Garcia-Alles, et al., Biochemistry 43:13037-13046, 2004).

[00171] The term "acetoin phosphate aminase" refers to a polypeptide having enzymatic activity that catalyzes the conversion of phosphoacetoin to 3-amino-2-butanol O-phosphate. Acetoin phosphate aminase may use the cofactor pyridoxal 5'-phosphate, NADH, or NADPH. The resulting product may have (R)- or (S)-stereochemistry at the 3-position. The pyridoxal phosphate-dependent enzyme may use an amino acid such as alanine or glutamate. The NADH-dependent and NADPH-dependent enzymes may use ammonia as a second substrate. Although there are no reports of enzymes catalyzing this reaction on phosphoacetoin, there is a pyridoxal phosphate-dependent enzyme that is proposed to carry out the analogous reaction on the similar substrate serinol phosphate (Yasuta, et al., Appl. Environ. Microbial. 67:4999-5009, 2001).

[00172] The term "aminobutanol phosphate phospholyase," also known as "amino alcohol O-phosphate lyase," refers to a polypeptide having enzymatic activity that catalyzes the conversion of 3-amino-2-butanol O-phosphate to 2-butanone. Amino butanol phosphate phospho-lyase may utilize the cofactor pyridoxal 5'-phosphate. There are reports of enzymes that catalyze the analogous reaction on the similar substrate 1-amino-2-propanol phosphate (Jones, et al., Biochem. J. 134:167-182, 1973). U.S. Patent Application Publication No. 2007/0259410 describes an aminobutanol phosphate phospho-lyase from the organism *Erwinia carotovora*.

[00173] The term "aminobutanol kinase" refers to a polypeptide having enzymatic activity that catalyzes the conversion of 3-amino-2-butanol to 3-amino-2-butanol O-phosphate. Amino butanol kinase may utilize ATP as the phosphate donor. Although there are no reports of enzymes catalyzing this reaction on 3-amino-2-butanol, there are reports of enzymes that catalyze the analogous reaction on the similar substrates ethanolamine and 1-amino-2-propanol (Jones, et al., *supra*). U.S. Patent Application Publication No.

2009/0155870 describes, in Example 14, an amino alcohol kinase of *Erwinia carotovora* subsp. *Atroseptica*.

[00174] The term "butanediol dehydrogenase," also known as "acetoin reductase," refers to a polypeptide having enzymatic activity that catalyzes the conversion of acetoin to 2,3-butanediol. Butanediol dehydrogenases are a subset of the broad family of alcohol dehydrogenases. Butanediol dehydrogenase enzymes may have specificity for production of (R)- or (S)-stereochemistry in the alcohol product. (S)-specific butanediol dehydrogenases are known as EC 1.1.1.76 and are available, for example, from *Klebsiella pneumoniae* (GenBank Nos: BBA13085, D86412). (R)-specific butanediol dehydrogenases are known as EC 1.1.1.4 and are available, for example, from *Bacillus cereus* (GenBank Nos. NP 830481, NC_004722; AAP07682, AE017000), and *Lactococcus lactis* (GenBank Nos. AAK04995, AE006323).

[00175] The term "butanediol dehydratase," also known as "dial dehydratase" or "propanediol dehydratase," refers to a polypeptide having enzymatic activity that catalyzes the conversion of 2,3-butanediol to 2-butanone. Butanediol dehydratase may utilize the cofactor adenosyl cobalamin (also known as coenzyme B₁₂ or vitamin B₁₂; although vitamin B₁₂ may refer also to other forms of cobalamin that are not coenzyme B₁₂). Adenosyl cobalamin-dependent enzymes are known as EC 4.2.1.28 and are available, for example, from *Klebsiella oxytoca* (GenBank Nos: AA08099 (alpha subunit), D45071; BAA08100 (beta subunit), D45071; and BBA08101 (gamma subunit), D45071; all three subunits are required for activity)), and *Klebsiella pneumonia* (GenBank Nos: AAC98384 (alpha subunit), AF102064; GenBank Nos: AAC98385 (beta subunit), AF102064, GenBank Nos: AAC98386 (gamma subunit), AF102064). Other suitable dial dehydratases include, but are not limited to, B₁₂-dependent dial dehydratases available from *Salmonella typhimurium* (GenBank Nos: AAB84102 (large subunit), AF026270; GenBank Nos: AAB84103 (medium subunit), AF026270; GenBank Nos: AAB84104 (small subunit), AF026270); and *Lactobacillus collinoides* (GenBank Nos: CAC82541 (large subunit), AJ297723; GenBank Nos: CAC82542 (medium subunit); AJ297723; GenBank Nos: CAD01091 (small subunit), AJ297723); and enzymes from *Lactobacillus brevis* (particularly strains CNRZ 734 and CNRZ 735, Speranza, et al., J. Agric. Food Chem. 45:3476-3480, 1997), and nucleotide sequences that encode the corresponding enzymes. Methods of dial dehydratase gene isolation are well known in the art (e.g., U.S. Patent No. 5,686,276).

[00176] In some embodiments, enzymes of the butanol biosynthetic pathway that are usually localized to the mitochondria are not localized to the mitochondria. In some embodiments, enzymes of the engineered butanol biosynthetic pathway may be localized to the cytosol. In some embodiments, an enzyme of the biosynthetic pathway may be localized to the cytosol by removing the mitochondrial targeting sequence. In some embodiments, mitochondrial targeting may be eliminated by generating new start codons as described, for example, in U.S. Patent No. 7,993,889, the entire contents of which are herein incorporated by reference. In some embodiments, the enzyme of the biosynthetic pathway that is localized to the cytosol is DHAD. In some embodiments, the enzyme from the biosynthetic pathway that is localized to the cytosol is KARI.

[00177] In some embodiments, the enzymes of the engineered butanol biosynthetic pathway may use NADH or NADPH as a co-factor, wherein NADH or NADPH acts as an electron donor. In some embodiments, one or more enzymes of the butanol biosynthetic pathway use NADH as an electron donor. In some embodiments, one or more enzymes of the butanol biosynthetic pathway use NADPH as an electron donor.

[00178] It will be appreciated that host cells comprising an isobutanol biosynthetic pathway as provided herein may further comprise one or more additional modifications. U.S. Patent Application Publication No. 2009/0305363, the entire contents of which are herein incorporated by reference, discloses increased conversion of pyruvate to acetolactate by engineering yeast for expression of a cytosol-localized acetolactate synthase and substantial elimination of pyruvate decarboxylase activity. In some embodiments, the host cells may comprise modifications to reduce glycerol-3-phosphate dehydrogenase activity and/or disruption in at least one gene encoding a polypeptide having pyruvate decarboxylase activity or a disruption in at least one gene encoding a regulatory element controlling pyruvate decarboxylase gene expression (as described in U.S. Patent Application Publication No. 2009/0305363, the entire contents of which are herein incorporated by reference), or modifications to a host cell that provide for increased carbon flux through an Entner-Doudoroff Pathway or reducing equivalents balance (as described in U.S. Patent Application Publication No. 2010/0120105, the entire contents of which are herein incorporated by reference). Other modifications include integration of at least one polynucleotide encoding a polypeptide that catalyzes a step in a pyruvate-utilizing biosynthetic pathway. Other modifications include at least one deletion, mutation, and/or substitution in an endogenous

polynucleotide encoding a polypeptide having acetolactate reductase activity. In some embodiments, the polypeptide having acetolactate reductase activity is YMR226C (SEQ ID NOs: 107, 108) of *Saccharomyces cerevisiae* or a homolog thereof. Additional modifications include a deletion, mutation, and/or substitution in an endogenous polynucleotide encoding a polypeptide having aldehyde dehydrogenase and/or aldehyde oxidase activity. In some embodiments, the polypeptide having aldehyde dehydrogenase activity is *ALD6* from *Saccharomyces cerevisiae* or a homolog thereof.

[00179] The term "pyruvate decarboxylase" refers to any polypeptide having a biological function of a pyruvate decarboxylase. Such polypeptides include a polypeptide that catalyzes the decarboxylation of pyruvic acid to acetaldehyde and carbon dioxide. Pyruvate dehydrogenases are known by the EC number 4.1.1.1. Such polypeptides can be determined by methods well known in the art and disclosed in U.S. Patent Application. Publication No. 2013/0071898, the entire contents of which are herein incorporated by reference. These enzymes are found in a number of yeast including *Saccharomyces cerevisiae* (GenBank Nos: CAA97575 (SEQ ID NO: 109), CAA97705 (SEQ ID NO: 111), CAA97091 (SEQ ID NO: 113)). Additional examples of PDC are provided in U.S. Patent Application. Publication No. 2009/035363, the entire contents of which are herein incorporated by reference.

[00180] A genetic modification which has the effect of reducing glucose repression wherein the yeast production host cell is *pdc-* is described in U.S. Patent Application Publication No. 2011/0124060, the entire contents of which are herein incorporated by reference. In some embodiments, the pyruvate decarboxylase that is deleted or down-regulated is selected from the group consisting of: *PDC1*, *PDC5*, *PDC6*, and combinations thereof. In some embodiments, the pyruvate decarboxylase is selected from those enzymes in Table 3. In some embodiments, host cells contain a deletion or down-regulation of a polynucleotide encoding a polypeptide that catalyzes the conversion of glyceraldehyde-3-phosphate to glycerate 1,3, bisphosphate. In some embodiments, the enzyme that catalyzes this reaction is glyceraldehyde-3-phosphate dehydrogenase.

Table 3. SEQ ID Numbers of PDC Target Gene coding regions and Proteins.

Description	SEQ ID NO: Amino Acid	SEQ ID NO: Nucleic Acid
PDC1 pyruvate decarboxylase from <i>Saccharomyces cerevisiae</i>	109	110
PDC5 pyruvate decarboxylase from <i>Saccharomyces cerevisiae</i>	111	112
PDC6 pyruvate decarboxylase <i>Saccharomyces cerevisiae</i>	113	114
pyruvate decarboxylase from <i>Candida glabrata</i>	115	116
PDC1 pyruvate decarboxylase from <i>Pichia stipitis</i>	117	118
PDC2 pyruvate decarboxylase from <i>Pichia stipitis</i>	119	120
pyruvate decarboxylase from <i>Kluyveromyces lactis</i>	121	122
pyruvate decarboxylase from <i>Yarrowia lipolytica</i>	123	124
pyruvate decarboxylase from <i>Schizosaccharomyces pombe</i>	125	126
pyruvate decarboxylase from <i>Zygosaccharomyces rouxii</i>	127	128

[00181] Yeasts may have one or more genes encoding pyruvate decarboxylase. For example, there is one gene encoding pyruvate decarboxylase in *Candida glabrata* and *Schizosaccharomyces pombe*, while there are three isozymes of pyruvate decarboxylase encoded by the *PDC1*, *PDC5*, and *PDC6* genes in *Saccharomyces*. In some embodiments, at least one PDC gene is inactivated. If the yeast cell used has more than one expressed (active) PDC gene, then each of the active PDC genes may be modified or inactivated thereby producing a pdc- cell. For example, in *Saccharomyces cerevisiae*, the *PDC1*, *PDC5*, and *PDC6* genes may be modified or inactivated. If a PDC gene is not active under the fermentation conditions to be used then such a gene would not need to be modified or inactivated.

[00182] Other target genes, such as those encoding pyruvate decarboxylase proteins having at least about 70-75%, at least about 75-85%, at least about 80-85%, at least about 85%-90%, at least about 90%-95%, or at least about 90%, or at least about 95%, or at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity to the pyruvate decarboxylases of SEQ ID NOs: 109, 111, 113, 115, 117, 119, 121, 123, 125, or 127 may be identified in the literature and in bioinformatics databases well known to the skilled person.

[00183] Recombinant host cells may further comprise (a) at least one heterologous polynucleotide encoding a polypeptide having dihydroxy-acid dehydratase activity; and (b)(i) at least one deletion, mutation, and/or substitution in an endogenous gene encoding a polypeptide affecting Fe-S cluster biosynthesis; and/or (ii) at least one heterologous polynucleotide encoding a polypeptide affecting Fe-S cluster biosynthesis. In some embodiments, the polypeptide affecting Fe-S cluster biosynthesis is encoded by *AFT1*, *AFT2*, *FRA2*, *GRX3* or *CCC1*. *AFT1* and *AFT2* are described by PCT Application Publication No. WO 2001/103300, the entire contents of which are herein incorporated by reference. In some embodiments, the polypeptide affecting Fe-S cluster biosynthesis is constitutive mutant *AFT1* L99A, *AFT1* L102A, *AFT1* C291F, or *AFT1* C293F.

Host Cells for Butanol Production

[00184] Recombinant microorganisms containing the genes necessary to encode the enzymatic pathway for conversion of a fermentable carbon substrate to butanol isomers may be constructed using techniques well known in the art. In the present invention, genes encoding the enzymes of one of the butanol biosynthetic pathways, for example, acetolactate synthase, acetohydroxy acid isomeroreductase, acetohydroxy acid dehydratase, branched-chain α -keto acid decarboxylase, and branched-chain alcohol dehydrogenase, may be isolated from various sources as described, for example, in U.S. Patent No. 7,993,889, the entire contents of which are herein incorporated by reference.

[00185] Once the relevant pathway genes are identified and isolated, the relevant enzymes of the butanol biosynthetic pathway may be introduced into the host cells or manipulated as described, for example, in U.S. Patent No. 7,993,889, the entire contents of which are herein incorporated by reference, to produce butanologens. The butanologens generated comprise

an engineered butanol biosynthetic pathway. In some embodiments, the butanologen is an isobutanologen, which comprises an engineered isobutanol biosynthetic pathway.

[00186] In some embodiments, the recombinant host cell may also comprise one or more polypeptides from a group of enzymes having the following Enzyme Commission Numbers: EC 2.2.1.6, EC 1.1.1.86, EC 4.2.1.9, EC 4.1.1.72, EC 1.1.1.1, EC 1.1.1.265, EC 1.1.1.2, EC 1.2.4.4, EC 1.3.99.2, EC 1.2.1.57, EC 1.2.1.10, EC 2.6.1.66, EC 2.6.1.42, EC 1.4.1.9, EC 1.4.1.8, EC 4.1.1.14, EC 2.6.1.18, EC 2.3.1.9, EC 2.3.1.16, EC 1.1.1.30, EC 1.1.1.35, EC 1.1.1.157, EC 1.1.1.36, EC 4.2.1.17, EC 4.2.1.55, EC 1.3.1.44, EC 1.3.1.38, EC 5.4.99.13, EC 4.1.1.5, EC 2.7.1.29, EC 1.1.1.76, EC 1.2.1.57, and EC 4.2.1.28.

[00187] In some embodiments, the recombinant host cell may comprise one or more polypeptides selected from acetolactate synthase, acetohydroxy acid isomeroreductase, acetohydroxy acid dehydratase, branched-chain alpha-keto acid decarboxylase, branched-chain alcohol dehydrogenase, acylating aldehyde dehydrogenase, branched-chain keto acid dehydrogenase, butyryl-CoA dehydrogenase, butyraldehyde dehydrogenase, transaminase, valine dehydrogenase, valine decarboxylase, omega transaminase, acetyl-CoA acetyltransferase, 3-hydroxybutyryl-CoA dehydrogenase, crotonase, butyryl-CoA dehydrogenase, isobutyryl-CoA mutase, acetolactate decarboxylase, acetoin aminase, butanol dehydrogenase, butyraldehyde dehydrogenase, acetoin kinase, acetoin phosphate aminase, aminobutanol phosphate phospholyase, aminobutanol kinase, butanediol dehydrogenase, and butanediol dehydratase.

[00188] In some embodiments, the recombinant host cell may be bacteria, cyanobacteria, filamentous fungi, or yeast. Suitable recombinant host cell capable of producing an alcohol (e.g., butanol) via a biosynthetic pathway include a member of the genera *Clostridium*, *Zymomonas*, *Escherichia*, *Salmonella*, *Serratia*, *Erwinia*, *Klebsiella*, *Shigella*, *Rhodococcus*, *Pseudomonas*, *Bacillus*, *Lactobacillus*, *Enterococcus*, *Alcaligenes*, *Klebsiella*, *Paenibacillus*, *Arthrobacter*, *Corynebacterium*, *Brevibacterium*, *Schizosaccharomyces*, *Kluyveromyces*, *Yarrowia*, *Pichia*, *Zygosaccharomyces*, *Debaryomyces*, *Candida*, *Brettanomyces*, *Pachysolen*, *Hansenula*, *Issatchenkia*, *Trichosporon*, *Yamadazyma*, or *Saccharomyces*. In some embodiments, the recombinant host cell may be selected from *Escherichia coli*, *Alcaligenes eutrophus*, *Bacillus licheniformis*, *Paenibacillus macerans*, *Rhodococcus erythropolis*, *Pseudomonas putida*, *Lactobacillus plantarum*, *Enterococcus faecium*, *Enterococcus gallinarum*, *Enterococcus faecalis*, *Bacillus subtilis*, *Candida sonorensis*,

Candida methanosorbosa, *Kluyveromyces lactis*, *Kluyveromyces marxianus*, *Kluyveromyces thermotolerans*, *Issatchenkia orientalis*, *Debaryomyces hansenii*, and *Saccharomyces cerevisiae*. In some embodiments, the recombinant host cell is yeast. In some embodiments, the recombinant host cell may be crabtree-positive yeast selected from *Saccharomyces*, *Zygosaccharomyces*, *Schizosaccharomyces*, *Dekkera*, *Torulopsis*, *Brettanomyces*, and some species of *Candida*. Species of crabtree-positive yeast include, but are not limited to, *Saccharomyces cerevisiae*, *Saccharomyces kluyveri*, *Schizosaccharomyces pombe*, *Saccharomyces bayanus*, *Saccharomyces mikitaie*, *Saccharomyces paradoxus*, *Saccharomyces uvarum*, *Saccharomyces castelli*, *Saccharomyces kluyveri*, *Zygosaccharomyces rouxii*, *Zygosaccharomyces bailli*, and *Candida glabrata*.

[00189] In some embodiments, the recombinant host cell may be a butanologen. In some embodiments, the butanologen may be an isobutanologen. In some embodiments, suitable isobutanologens include any yeast host useful for genetic modification and recombinant gene expression. In some embodiments, the host cell is a member of the genera *Saccharomyces*. In some embodiments, the host cell is *Saccharomyces cerevisiae*. *Saccharomyces cerevisiae* yeast are known in the art and are available from a variety of sources including, but not limited to, American Type Culture Collection (Rockville, MD), Centraalbureau voor Schimmelcultures (CBS) Fungal Biodiversity Centre, LeSaffre, Gert Strand AB, Ferm Solutions, North American Bioproducts, Martrex, and Lallemand. *Saccharomyces cerevisiae* include, but are not limited to, BY4741, CEN.PK 113-7D, Ethanol Red® yeast, Ferm Pro™ yeast, Bio-Ferm® XR yeast, Gert Strand Prestige Batch Turbo alcohol yeast, Gert Strand Pot Distillers yeast, Gert Strand Distillers Turbo yeast, FerMax™ Green yeast, FerMax™ Gold yeast, Thermosacc® yeast, BG-1, PE-2, CAT-1, CBS7959, CBS7960, and CBS7961.

[00190] In some embodiments, the butanologen expresses an engineered butanol biosynthetic pathway. In some embodiments, the butanologen is an isobutanologen expressing an engineered isobutanol biosynthetic pathway.

[00191] In some embodiments, the engineered isobutanol pathway comprises the following substrate to product conversions:

- a) pyruvate to acetolactate
- b) acetolactate to 2,3-dihydroxyisovalerate
- c) 2,3-dihydroxyisovalerate to α -ketoisovalerate
- d) α -ketoisovalerate to isobutyraldehyde, and

e) isobutyraldehyde to isobutanol.

[00192] In some embodiments, one or more of the substrate to product conversions utilizes NADH or NADPH as a cofactor.

[00193] In some embodiments, enzymes from the biosynthetic pathway may be localized to the cytosol. In some embodiments, enzymes from the biosynthetic pathway that are usually localized to the mitochondria may be localized to the cytosol. In some embodiments, an enzyme from the biosynthetic pathway may be localized to the cytosol by removing the mitochondrial targeting sequence. In some embodiments, mitochondrial targeting may be eliminated by generating new start codons as described in, for example, U.S. Patent No. 7,851,188, the entire contents of which are herein incorporated by reference. In some embodiments, the enzyme from the biosynthetic pathway that is localized to the cytosol is DHAD. In some embodiments, the enzyme from the biosynthetic pathway that is localized to the cytosol is KARI.

Production of Butanol

[00194] Disclosed herein are processes suitable for production of butanol from a carbon substrate and employing a recombinant host cell. In some embodiments, recombinant host cells may comprise an isobutanol biosynthetic pathway such as, but not limited to, isobutanol biosynthetic pathways disclosed herein. The ability to utilize carbon substrates to produce isobutanol can be confirmed using methods known in the art including, but not limited to, those described in U.S. Patent No. 7,851,188, the entire contents of which are herein incorporated by reference. For example, to confirm utilization of sucrose to produce isobutanol, the concentration of isobutanol in the culture media can be determined by a number of methods known in the art. For example, a specific high performance liquid chromatography (HPLC) method utilized a Shodex SH-1011 column with a Shodex SH-G guard column (Waters Corporation, Milford, MA), with refractive index (RI) detection. Chromatographic separation was achieved using 0.01 M H₂SO₄ as the mobile phase with a flow rate of 0.5 mL/min and a column temperature of 50°C. Isobutanol had a retention time of 46.6 min under the conditions used. Alternatively, gas chromatography (GC) methods are available. For example, a specific GC method utilized an HP-INNOWax column (30 m x 0.53 mm id, 1 µm film thickness, Agilent Technologies, Wilmington, DE), with a flame ionization detector (FID). The carrier gas was helium at a flow rate of 4.5 mL/min,

measured at 150°C with constant head pressure; injector split was 1:25 at 200°C; oven temperature was 45°C for 1 min, 45 to 220°C at 10°C/min, and 220°C for 5 min; and FID detection was employed at 240°C with 26 mL/min helium makeup gas. The retention time of isobutanol was 4.5 min.

Carbon substrates

- [00195] Suitable carbon substrates may include, but are not limited to, monosaccharides such as fructose or glucose; oligosaccharides such as lactose, maltose, galactose, or sucrose; polysaccharides such as starch; cellulose; or mixtures thereof, and unpurified mixtures from renewable feedstocks such as cheese whey permeate, cornsteep liquor, sugar beet molasses, and barley malt. Other carbon substrates may include ethanol, lactate, succinate, or glycerol.
- [00196] In some embodiments, the carbon substrate may be oligosaccharides, polysaccharides, monosaccharides, and mixtures thereof. In some embodiments, the carbon substrate may be fructose, glucose, lactose, maltose, galactose, sucrose, starch, cellulose, feedstocks, ethanol, lactate, succinate, glycerol, corn mash, sugar cane, a C5 sugar such as xylose and arabinose, and mixtures thereof.
- [00197] Additionally, the carbon substrate may also be one-carbon substrates such as carbon dioxide or methanol for which metabolic conversion into key biochemical intermediates has been demonstrated. In addition to one and two carbon substrates, methylotrophic organisms are also known to utilize a number of other carbon containing compounds such as methylamine, glucosamine and a variety of amino acids for metabolic activity. For example, methylotrophic yeasts are known to utilize the carbon from methylamine to form trehalose or glycerol (Bellion, et al., *Microb. Growth C1 Compd.*, [Int. Symp.], 7th (1993), 415-32, Editor(s): Murrell, J. Collin; Kelly, Don P. Publisher: Intercept, Andover, UK). Similarly, various species of *Candida* will metabolize alanine or oleic acid (Sulter, et al., *Arch. Microbiol.* 153:485-489, 1990). Hence, it is contemplated that the source of carbon utilized in the present invention may encompass a wide variety of carbon containing substrates and will only be limited by the choice of organism.
- [00198] Although it is contemplated that all of the above mentioned carbon substrates and mixtures thereof are suitable in the present invention, in some embodiments, the carbon substrates are glucose, fructose, and sucrose, or mixtures of these with C5 sugars such as xylose and arabinose for yeasts cells modified to use C5 sugars. Sucrose may be derived

from renewable sugar sources such as sugar cane, sugar beets, cassava, sweet sorghum, and mixtures thereof. Glucose and dextrose may be derived from renewable grain sources through saccharification of starch based feedstocks including grains such as corn, wheat, rye, barley, oats, and mixtures thereof. In addition, fermentable sugars may be derived from renewable cellulosic or lignocellulosic feedstock through processes of pretreatment and saccharification as described, for example, in U.S. Patent Application Publication No. 2007/0031918, the entire contents of which are herein incorporated by reference. Feedstock includes materials comprising cellulose, and optionally further comprising hemicellulose, lignin, starch, oligosaccharides, and/or monosaccharides. Feedstock may also comprise additional components, such as protein and/or lipid. Feedstock may be derived from a single source, or feedstock can comprise a mixture derived from more than one source; for example, feedstock may comprise a mixture of corn cobs and corn stover, or a mixture of grass and leaves. Feedstock includes, but is not limited to, bioenergy crops, agricultural residues, municipal solid waste, industrial solid waste, sludge from paper manufacture, yard waste, wood and forestry waste. Examples of feedstock include, but are not limited to, corn grain, corn cobs, crop residues such as corn husks, corn stover, grasses, wheat, wheat straw, barley, barley straw, hay, rice straw, switchgrass, waste paper, sugar cane bagasse, sorghum, soy, components obtained from milling of grains, trees, branches, roots, leaves, wood chips, sawdust, shrubs and bushes, vegetables, fruits, flowers, animal manure, and mixtures thereof. Methods for preparing feedstock are described in U.S. Patent Application Publication No. 2012/0164302, the entire contents of which are herein incorporated by reference. In some embodiments, the carbon substrate is glucose derived from corn. In some embodiments, the carbon substrate is glucose derived from wheat. In some embodiments, the carbon substrate is sucrose derived from sugar cane.

[00199] In some embodiments, the recombinant host cell is contacted with carbon substrates under conditions whereby isobutanol is produced. In some embodiments, the recombinant host cell at a given cell density may be added to a fermentation vessel along with suitable media. In some embodiments, the media may contain the carbon substrate, or the carbon substrate may be added separately. In some embodiments, the carbon substrate may be present at any concentration at the start of and/or during production of isobutanol. In some embodiments, the initial concentration of carbon substrate may be in the range of about 60 to 80 g/L. Suitable temperatures for fermentation are known to those of skill in the art and will

depend on the genus and/or species of the recombinant host cell employed. In some embodiments, suitable temperatures are in the range of 25°C to 43°C. The contact between the recombinant host cell and the carbon substrate may be any length of time whereby isobutanol is produced. In some embodiments, the contact occurs for at least about 8 hours, at least about 24 hours, at least about 48 hours. In some embodiments, the contact occurs for less than 8 hours. In some embodiments, the contact occurs until at least about 90% of the carbon substrate is utilized or until a desired effective titer of isobutanol is reached. In some embodiments, the effective titer of isobutanol is at least about 40 g/L, at least about 50 g/L, at least about 60 g/L, at least about 70 g/L, at least about 80 g/L, at least about 90 g/L, at least about 100 g/L, or at least about 110 g/L.

[00200] In some embodiments, the recombinant host cell produces butanol at least about 90% of effective yield, at least about 91% of effective yield, at least about 92% of effective yield, at least about 93% of effective yield, at least about 94% of effective yield, at least about 95% of effective yield, at least about 96% of effective yield, at least about 97% of effective yield, at least about 98% of effective yield, or at least about 99% of effective yield. In some embodiments, the recombinant host cell produces butanol at least about 55% to at least about 75% of effective yield, at least about 50% to at least about 80% of effective yield, at least about 45% to at least about 85% of effective yield, at least about 40% to at least about 90% of effective yield, at least about 35% to at least about 95% of effective yield, at least about 30% to at least about 99% of effective yield, at least about 25% to at least about 99% of effective yield, at least about 10% to at least about 99% of effective yield or at least about 10% to at least about 100% of effective yield.

[00201] In some embodiments, the recombinant host cell may be incubated at a temperature range of 30°C to 37°C. In some embodiments, the recombinant host cell may be incubated at for a time period of one to five hours. In some embodiments, the recombinant host cell may be incubated with agitation (e.g., 100 to 400 rpm) in shakers (Innova 44R, New Brunswick Scientific, CT).

[00202] In some embodiments, the recombinant host cell is present at a cell density of at least about 0.5 gdcw/L at the first contacting with the carbon substrate. In some embodiments, the recombinant host cell may be grown to a cell density of at least about 6 gdcw/L prior to contacting with carbon substrate for the production of isobutanol. In some embodiments, the cell density may be at least about 20 gdcw/L, at least about 25 gdcw/L, or

at least about 35 gdcw/L, prior to contact with carbon substrate. In some embodiments, the recombinant host cell is present at a cell density of at least about 6 gdcw/L to 30 gdcw/L during the first contacting with the carbon substrate. In some embodiments, the cell density of the recombinant host cell may be 6.5 gdcw/L, 7 gdcw/L, 7.5 gdcw/L, 8 gdcw/L, 8.5 gdcw/L, 9 gdcw/L, 9.5 gdcw/L, 10 gdcw/L, 10.5 gdcw/L, 12 gdcw/L, 15 gdcw/L, 17 gdcw/L, 20 gdcw/L, 22 gdcw/L, 25 gdcw/L, 27 gdcw/L, or 30 gdcw/L during the first contacting with the carbon substrate.

[00203] In some embodiments, the recombinant host cell has a specific productivity of at least about 0.1 g/gdcw/h. In some embodiments, butanol is produced at an effective rate of at least about 0.1 g/gdcw/h during the first contacting with the carbon substrate. In some embodiments, the first contacting with the carbon substrate occurs in the presence of an extractant. In some embodiments, the recombinant host cell maintains a sugar uptake rate of at least about 1.0 g/gdcw/h. In some embodiments, the recombinant host cell maintains a sugar uptake rate of at least about 0.5 g/g/hr. In some embodiments, the glucose utilization rate is at least about 2.5 g/gdcw/h. In some embodiments, the sucrose uptake rate is at least about 2.5 g/gdcw/h. In some embodiments, the combined glucose and fructose uptake rate is at least about 2.5 g/gdcw/h. In some embodiments, the first contacting with the carbon substrate occurs in anaerobic conditions. In some embodiments, the first contacting with the carbon substrate occurs in microaerobic conditions. In some embodiments, cell recycling occurs in anaerobic conditions. In some embodiments, cell recycling occurs in microaerobic conditions.

Fermentation Conditions

[00204] Cells may be grown at a temperature in the range of about 20°C to about 40°C in an appropriate medium. In some embodiments, the cells are grown at a temperature of 20°C, 22°C, 25°C, 27°C, 30°C, 32°C, 35°C, 37°C, or 40°C. Suitable growth media in the present invention include common commercially prepared media such as Sabouraud Dextrose (SD) broth, Yeast Medium (YM) broth, or broth that includes yeast nitrogen base, ammonium sulfate, and dextrose (as the carbon/energy source) or YPD Medium, a blend of peptone, yeast extract, and dextrose in optimal proportions for growing most *Saccharomyces cerevisiae* strains. Other defined or synthetic growth media may also be used, and the appropriate medium for growth of the particular microorganism will be known by one skilled

in the art of microbiology or fermentation science. The use of agents known to modulate catabolite repression directly or indirectly, for example, cyclic adenosine 2':3'-monophosphate, may also be incorporated into the fermentation medium.

[00205] In addition to an appropriate carbon source, fermentation media may contain minerals, vitamins, amino acids (e.g., glycine, proline), salts, cofactors, unsaturated fats, steroids, buffers, and other components, known to those skilled in the art, suitable for the growth of the cultures and promotion of an enzymatic pathway described herein. For example, the medium may contain one or more of the following: biotin, pantothenate, folic acid, niacin, aminobenzoic acid, pyridoxine, riboflavin, thiamine, inositol, potassium (e.g., potassium phosphate), boric acid, calcium (e.g., calcium chloride), chromium, copper (e.g., copper sulfate), iodide (e.g., potassium iodide), iron (e.g., ferric chloride), lithium, magnesium (e.g., magnesium sulfate, magnesium chloride), manganese (e.g., manganese sulfate), molybdenum, calcium chloride, sodium chloride, silicon, vanadium, zinc (e.g., zinc sulfate), yeast extract, soy peptone, and the like.

[00206] In some embodiments of the present invention, the fermentation medium may comprise magnesium in the range of about 5 mM to about 250 mM. In some embodiments, the fermentation medium may comprise magnesium in the range of about 5 mM to about 200 mM. In some embodiments, the fermentation medium may comprise magnesium in the range of about 10 mM to about 200 mM. In some embodiments, the fermentation medium may comprise magnesium in the range of about 50 mM to about 200 mM. In some embodiments, the fermentation medium may comprise magnesium in the range of about 100 mM to about 200 mM. In some embodiments, the fermentation medium may comprise magnesium in the range of about 10 mM to about 150 mM. In some embodiments, the fermentation medium may comprise magnesium in the range of about 50 mM to about 150 mM. In some embodiments, the fermentation medium may comprise magnesium in the range of about 100 mM to about 150 mM. In some embodiments, the fermentation medium may comprise magnesium in the range of about 30 mM to about 100 mM. In some embodiments, the fermentation medium may comprise magnesium in the range of about 30 mM to about 70 mM.

[00207] In some embodiments, the amount of magnesium in the fermentation medium is about 5 mM, about 10 mM, about 15 mM, about 20 mM, about 25 mM, about 30 mM, about 35 mM, about 40 mM, about 45 mM, about 50 mM, about 55 mM, about 60 mM, about

65 mM, about 70 mM, about 75 mM, about 80 mM, about 85 mM, about 90 mM, about 95 mM, about 100 mM, about 105 mM, about 110 mM, about 115 mM, about 120 mM, about 125 mM, about 130 mM, about 135mM, about 140 mM, about 145mM, about 150 mM, about 155mM, about 160 mM, about 165mM, about 170 mM, about 175mM, about 180 mM, about 185mM, about 190 mM, about 195mM, about 200 mM, about 205 mM, about 210 mM, about 215 mM, about 220 mM, about 225 mM, about 230 mM, about 235 mM, about 240 mM, about 245 mM, or about 250 mM. In some embodiments, the fermentation medium may be supplemented with magnesium chloride, magnesium sulfate, other magnesium salts, or mixtures thereof.

[00208] In some embodiments, magnesium may be added during preparation of the feedstock or biomass. In some embodiments, magnesium may be added during the fermentation process. In some embodiments, magnesium in the range of about 5 mM to about 250 mM may be maintained in the fermentation medium during the fermentation process. In some embodiments, magnesium in the range of about 5 mM to about 200 mM may be maintained in the fermentation medium during the fermentation process. In some embodiments, magnesium in the range of about 10 mM to about 200 mM may be maintained in the fermentation medium during the fermentation process. In some embodiments, magnesium in the range of about 50 mM to about 200 mM may be maintained in the fermentation medium during the fermentation process. In some embodiments, magnesium in the range of about 100 mM to about 200 mM may be maintained in the fermentation medium during the fermentation process. In some embodiments, magnesium in the range of about 10 mM to about 150 mM may be maintained in the fermentation medium during the fermentation process. In some embodiments, magnesium in the range of about 50 mM to about 150 mM may be maintained in the fermentation medium during the fermentation process. In some embodiments, magnesium in the range of about 100 mM to about 150 mM may be maintained in the fermentation medium during the fermentation process. In some embodiments, magnesium in the range of about 30 mM to about 100 mM may be maintained in the fermentation medium during the fermentation process. In some embodiments, magnesium in the range of about 30 mM to about 70 mM may be maintained in the fermentation medium during the fermentation process.

[00209] In some embodiments, it may be beneficial to maintain low calcium-to-magnesium ratio in the fermentation medium. In some embodiments, calcium may be removed from the

fermentation medium by precipitation or ion exchange chromatography. In some embodiments, the concentrations of calcium may be managed by supplementing the fermentation medium with magnesium.

[00210] In some embodiments, nutrients such as minerals, vitamins, amino acids, trace elements, and other components (e.g., calcium, iron, potassium, magnesium, manganese, sodium, phosphorus, sulfur, and zinc) may be provided by the supplementation of the feedstock, feedstock preparation, or fermentation broth with backset. In some embodiments, feedstock, feedstock preparation, and/or fermentation broth may be supplemented with about 10% to about 100% of backset (e.g., percentage of total backset generated by processing of whole stillage). In some embodiments, about 10%, about 15%, about 20%, about 25%, about 30%, about 35%, about 40%, about 45%, about 50%, about 60%, about 70%, about 80%, about 90%, about 95%, or 100% of backset (e.g., percentage of total backset generated by processing of whole stillage) may be used to supplement feedstock, feedstock preparation, and/or fermentation broth.

[00211] In some embodiments, backset may be added to feedstock, feedstock preparation, and/or fermentation broth as a percentage of the water volume of feedstock, feedstock preparation, and/or fermentation broth. In some embodiments, backset may be added as about 5%, about 10%, about 15%, about 20%, about 25%, about 30%, about 35%, about 40%, about 45%, or about 50% of the water volume of feedstock, feedstock preparation, and/or or fermentation broth.

[00212] In some embodiments, the fermentation medium may further contain butanol. In some embodiments, the butanol is in the range of about 0.01 mM to about 500 mM. In some embodiments, the butanol is about 0.01 mM, about 1.0 mM, about 10 mM, about 15 mM, about 20 mM, about 25 mM, about 30 mM, about 35 mM, about 40 mM, about 45 mM, about 50 mM, about 55 mM, about 60 mM, about 65 mM, about 70 mM, about 75 mM, about 80 mM, about 85 mM, about 90 mM, about 95 mM, about 100 mM, about 110 mM, about 120 mM, about 130 mM, about 140 mM, about 150 mM, about 160 mM, about 170 mM, about 180 mM, about 190 mM, about 200 mM, about 210 mM, about 220 mM, about 230 mM, about 240 mM, about 250 mM, about 260 mM, about 270 mM, about 280 mM, about 290 mM, about 300 mM, about 310 mM, about 320 mM, about 330 mM, about 340 mM, about 350 mM, about 360 mM, about 370 mM, about 380 mM, about 390 mM, about 400 mM, about 410 mM, about 420 mM, about 430 mM, about 440 mM,

about 450 mM, about 460 mM, about 470 mM, about 480 mM, about 490 mM or about 500 mM. In some embodiments, butanol present in the fermentation medium is from about 0.01% to about 100% of the theoretical yield of butanol. In some embodiments, butanol present in the fermentation medium is 0.01%, 0.5%, 1%, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% or 100% of the theoretical yield of butanol.

[00213] Suitable pH ranges for the fermentation are from about pH 3.0 to about pH 9.0. In some embodiments, about pH 4.0 to about pH 8.0 may be used for the initial condition. In some embodiments, about pH 5.0 to about pH 9.0 may be used for the initial condition. In some embodiments, about pH 3.5 to about pH 9.0 may be used for the initial condition. In some embodiments, about pH 4.5 to about pH 6.5 may be used for the initial condition. In some embodiments, about pH 5.0 to about pH 8.0 may be used for the initial condition. In some embodiments, about pH 6.0 to about pH 8.0 may be used for the initial condition. Suitable pH ranges for the fermentation of yeast are typically from about pH 3.0 to about pH 9.0. Suitable pH ranges for the fermentation of other microorganisms are from about pH 3.0 to about pH 7.5.

[00214] Fermentations may be performed under aerobic or anaerobic conditions. In some embodiments, anaerobic or microaerobic conditions are used for fermentations.

[00215] In some embodiments, butanol may be produced in one or more of the following growth phases: high growth log phase, moderate through static lag phase, stationary phase, steady state growth phase, and combinations thereof.

[00216] In some embodiments, the recombinant host cell may be propagated in a propagation tank. In some embodiments, the recombinant host cell from the propagation tank may be used to inoculate one or more fermentors. In some embodiments, the propagation tank may comprise one or more of the following mash, water, enzymes, nutrients, and microorganisms. In some embodiments, magnesium may be added to the propagation tank. In some embodiments, the recombinant host cell may be pre-conditioned by the addition of magnesium.

Industrial Batch and Continuous Fermentations

[00217] In some embodiments, butanol or butanol isomers may be produced using batch or continuous fermentation. Butanol isomers such as isobutanol may be produced using a batch

method of fermentation. A classical batch fermentation is a closed system where the composition of the medium is set at the beginning of the fermentation and not subject to artificial alterations during the fermentation. For example, at the beginning of the fermentation, the medium is inoculated with the desired organism or organisms, and fermentation is permitted to occur without adding anything to the system. Typically, a "batch" fermentation is batch with respect to the addition of carbon source and attempts are often made at controlling factors such as pH and oxygen concentration. In batch systems, the metabolite and biomass compositions of the system change constantly up to the time the fermentation is stopped. Within batch cultures, cells moderate through a static lag phase to a high growth log phase and finally to a stationary phase where growth rate is diminished or halted. If untreated, cells in the stationary phase will eventually die. Cells in log phase generally are responsible for the bulk of production of end product or intermediate.

[00218] A variation on the standard batch system is the fed-batch system. Fed-batch fermentation processes are also suitable in the present invention and may comprise a batch system with the exception that the substrate is added in increments as the fermentation progresses. Fed-batch systems are useful when catabolite repression is apt to inhibit the metabolism of the cells and where it is desirable to have limited amounts of substrate in the media. Batch and fed-batch fermentations are common and well known in the art and examples may be found in Thomas D. Brock in *Biotechnology: A Textbook of Industrial Microbiology*, Second Edition (1989) Sinauer Associates, Inc., Sunderland, MA., or Deshpande, *Appl. Biochem. Biotechnol.* 36:227, 1992.

[00219] Butanol may also be produced using continuous fermentation methods. Continuous fermentation is an open system where a defined fermentation medium is added continuously to a bioreactor and an equal amount of conditioned media is removed simultaneously for processing. Continuous fermentation generally maintains the cultures at a constant high density where cells are primarily in log phase growth. Continuous fermentation allows for the modulation of one factor or any number of factors that affect cell growth or end product concentration. Methods of modulating nutrients and growth factors for continuous fermentation processes as well as techniques for maximizing the rate of product formation are well known in the art of industrial microbiology and a variety of methods are detailed by Brock, *supra*.

[00220] It is contemplated that the production of isobutanol, or other products, may be practiced using batch, fed-batch or continuous processes and that any known mode of fermentation would be suitable. Additionally, it is contemplated that cells may be immobilized on a substrate as whole cell catalysts and subjected to fermentation conditions for isobutanol production.

Methods for Butanol Isolation from the Fermentation Medium

[00221] Bioproducted butanol or butanol isomers such as isobutanol may be isolated from the fermentation medium using methods known in the art for ABE fermentations (see, e.g., Durre, Appl. Microbiol. Biotechnol. 49:639-648, 1998; Groot, et al., Process. Biochem. 27:61-75, 1992, and references therein). For example, solids may be removed from the fermentation medium by centrifugation, filtration, decantation, or the like. Then, the isobutanol may be isolated from the fermentation medium using methods such as distillation, azeotropic distillation, liquid-liquid extraction, adsorption, gas stripping, membrane evaporation, or pervaporation.

[00222] Because isobutanol forms a low boiling point, azeotropic mixture with water, distillation can be used to separate the mixture up to its azeotropic composition. Distillation may be used in combination with another separation method to obtain separation around the azeotrope. Methods that may be used in combination with distillation to isolate and purify isobutanol include, but are not limited to, decantation, liquid-liquid extraction, adsorption, and membrane-based techniques. Additionally, isobutanol may be isolated using azeotropic distillation using an entrainer (see, e.g., Doherty and Malone, *Conceptual Design of Distillation Systems*, McGraw Hill, New York, 2001).

[00223] The isobutanol-water mixture forms a heterogeneous azeotrope so that distillation may be used in combination with decantation to isolate and purify the isobutanol. In this method, the isobutanol containing fermentation broth is distilled to near the azeotropic composition. Then, the azeotropic mixture is condensed, and the isobutanol is separated from the fermentation medium by decantation. The decanted aqueous phase may be returned to the first distillation column as reflux. The isobutanol-rich decanted organic phase may be further purified by distillation in a second distillation column.

[00224] The isobutanol can also be isolated from the fermentation medium using liquid-liquid extraction in combination with distillation. In this method, the isobutanol is extracted

from the fermentation broth using liquid-liquid extraction with a suitable solvent. The isobutanol-containing organic phase is then distilled to separate the isobutanol from the solvent.

[00225] Distillation in combination with adsorption can also be used to isolate isobutanol from the fermentation medium. In this method, the fermentation broth containing the isobutanol is distilled to near the azeotropic composition and then the remaining water is removed by use of an adsorbent such as molecular sieves (Aden, et al., *Lignocellulosic Biomass to Ethanol Process Design and Economics Utilizing Co-Current Dilute Acid Prehydrolysis and Enzymatic Hydrolysis for Corn Stover*, Report NREL/TP-510-32438, National Renewable Energy Laboratory, June 2002).

[00226] Additionally, distillation in combination with pervaporation may be used to isolate and purify isobutanol from the fermentation medium. In this method, the fermentation broth containing the isobutanol is distilled to near the azeotropic composition, and then the remaining water is removed by pervaporation through a hydrophilic membrane (Guo, et al., *J. Membr. Sci.* 245:199-210, 2004).

[00227] *In situ* product removal (ISPR) (also referred to as extractive fermentation) can be used to remove isobutanol (or other fermentative alcohol) from the fermentation vessel as it is produced, thereby allowing the microorganism to produce isobutanol at high yields. One method for ISPR for removing fermentative alcohol that has been described in the art is liquid-liquid extraction. In general, with regard to isobutanol fermentation, for example, the fermentation medium, which includes the microorganism, is contacted with an organic extractant at a time before the isobutanol concentration reaches a toxic level. The organic extractant and the fermentation medium form a biphasic mixture. The isobutanol partitions into the organic extractant phase, decreasing the concentration in the aqueous phase containing the microorganism, thereby limiting the exposure of the microorganism to the inhibitory isobutanol.

[00228] Liquid-liquid extraction can be performed, for example, according to the processes described in U.S. Patent Application Publication No. 2009/0305370, the entire contents of which are herein incorporated by reference. U.S. Patent Application Publication No. 2009/0305370 describes methods for producing and recovering isobutanol from a fermentation broth using liquid-liquid extraction, the methods comprising the step of contacting the fermentation broth with a water immiscible extractant to form a two-phase

mixture comprising an aqueous phase and an organic phase. Extractant may be one or more organic extractants such as saturated, mono-unsaturated, poly-unsaturated (and mixtures thereof) C_{12} to C_{22} fatty alcohols, C_{12} to C_{22} fatty acids, esters of C_{12} to C_{22} fatty acids, C_{12} to C_{22} fatty aldehydes, and mixtures thereof. The extractants may also be non-alcohol extractants. The extractants may be an exogenous organic extractant such as oleyl alcohol, behenyl alcohol, cetyl alcohol, lauryl alcohol, myristyl alcohol, stearyl alcohol, alkyl alkanols, 1-undecanol, oleic acid, lauric acid, myristic acid, stearic acid, methyl myristate, methyl oleate, undecanal, lauric aldehyde, 20-methylundecanal, trioctyl phosphine oxide, and mixtures thereof. In some embodiments, the extractant may be corn oil fatty acids.

[00229] In some embodiments, an ester can be formed by contacting the alcohol in a fermentation medium with an organic acid (e.g., fatty acids) and a catalyst capable of esterifying the alcohol with the organic acid. In such embodiments, the organic acid can serve as an ISPR extractant into which the alcohol esters partition. The organic acid can be supplied to the fermentation vessel and/or derived from the feedstock supplying fermentable carbon fed to the fermentation vessel. Lipids present in the feedstock can be catalytically hydrolyzed to organic acid, and the same catalyst (e.g., enzymes) can esterify the organic acid with the alcohol. The catalyst can be supplied to the feedstock prior to fermentation, or can be supplied to the fermentation vessel before or contemporaneously with the supplying of the feedstock. When the catalyst is supplied to the fermentation vessel, alcohol esters can be obtained by hydrolysis of the lipids into organic acid and substantially simultaneous esterification of the organic acid with the alcohol present in the fermentation vessel. Organic acid and/or native oil not derived from the feedstock can also be fed to the fermentation vessel, with the native oil being hydrolyzed into organic acid. Any organic acid not esterified with the alcohol can serve as part of the ISPR extractant. The extractant containing alcohol esters can be separated from the fermentation medium, and the alcohol can be recovered from the extractant. The extractant can be recycled to the fermentation vessel. Thus, in the case of isobutanol production, for example, the conversion of isobutanol to an ester reduces the free isobutanol concentration in the fermentation medium, shielding the microorganism from the toxic effect of increasing isobutanol concentration. In addition, unfractionated grain can be used as feedstock without separation of lipids therein, since the lipids can be catalytically hydrolyzed to organic acid, thereby decreasing the rate of build-up of lipids in the ISPR extractant. Other isobutanol product recovery and/or ISPR methods

may be employed including those described in U.S. Patent Application Publication No. 2011/0097773, U.S. Patent Application Publication No. 2011/0159558, U.S. Patent Application Publication No. 2011/0136193, and U.S. Patent Application Publication No. 2012/0156738, the entire contents of each are herein incorporated by reference.

[00230] *In situ* product removal can be carried out in a batch mode or a continuous mode. In a continuous mode of *in situ* product removal, product is continually removed from the reactor. In a batchwise mode of *in situ* product removal, an organic extractant is added to the fermentation vessel and the extractant is not removed during the process. For *in situ* product removal, the organic extractant can contact the fermentation medium at the start of the fermentation forming a biphasic fermentation medium. Alternatively, the organic extractant can contact the fermentation medium after the microorganism has achieved a desired amount of growth, which can be determined by measuring the optical density of the culture. Further, the organic extractant can contact the fermentation medium at a time at which the alcohol level in the fermentation medium reaches a preselected level. In the case of isobutanol production according to some embodiments of the present invention, the organic extractant can contact the fermentation medium at a time before the isobutanol concentration reaches a toxic level, so as to esterify the isobutanol with the organic acid to produce isobutanol esters and consequently reduce the concentration of isobutanol in the fermentation vessel. The ester-containing organic phase can then be removed from the fermentation vessel (and separated from the fermentation broth which constitutes the aqueous phase) after a desired effective titer of the isobutanol esters is achieved. In some embodiments, the ester-containing organic phase is separated from the aqueous phase after fermentation of the available fermentable sugar in the fermentation vessel is substantially complete.

[00231] Isobutanol titer in any phase can be determined by methods known in the art such as via high performance liquid chromatography (HPLC) or gas chromatography (GC), as described, for example, in U.S. Patent Application Publication No. 2009/0305370, the entire contents of which are herein incorporated by reference.

[00232] Following fermentation, the fermentation medium may be further processed to produce dried distillers grains and solubles (DDGS) and thin stillage. For example, the fermentation medium may be transferred to a beer column generating an alcohol-rich vaporized stream, which may be processed for the recovery of the alcohol, and a bottoms stream known as whole stillage. Whole stillage contains unfermented solids (e.g., distiller's

- 74 -

grain solids), dissolved materials (e.g., carbon substrates, minerals, vitamins, amino acids, trace elements, and other components), and water. Whole stillage may be processed using any known separation technique including centrifugation, filtration, screen separation, hydroclone, or any other means for separating liquids from solids. Separation of whole stillage generates a solids stream (e.g., wet cake) and a liquid stream known as thin stillage. Thin stillage may be further processed for water removal, for example, by evaporation. Examples of evaporation systems are described in U.S. Patent Application Publication No. 2011/0315541, the entire contents of which are herein incorporated by reference. Evaporation incrementally evaporates water from the thin stillage to eventually produce a syrup, which may be combined with the wet cake to yield DDGS.

[00233] Thin stillage may also be used in feedstock preparation as a replacement for water (known as "backsetting"). Using backset as a replacement for water can result in reduced capitol and energy costs. In addition, as thin stillage ("backset") comprises dissolved materials such as carbon substrates, minerals, vitamins, amino acids, trace elements, and other components, thin stillage or backset may also be used as a source of nutrient supplementation for fermentation. As such, the additional nutrient supplementation may improve biomass growth, fermentation rate, and tolerance.

[00234] All documents cited herein, including journal articles or abstracts, published or corresponding U.S. or foreign patent applications, issued or foreign patents, or any other documents, are each entirely incorporated by reference herein, including all data, tables, figures, and text presented in the cited documents.

[00235] Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.

EXAMPLES

[00236] The present invention is further defined in the following Examples. It should be understood that these Examples, while indicating embodiments of the invention, are given by way of illustration only. From the above discussion and these Examples, one skilled in the art can ascertain the essential characteristics of this invention, and without departing from the spirit and scope thereof, can make various changes and modifications of the invention to adapt it to various uses and conditions.

[00237] The meaning of abbreviations is as follows: "sec" means second(s), "min" means minute(s), "h" means hour(s), "nm" means nanometer(s), "mm" means millimeter(s), "uL" means microliter(s), "mL" means milliliter(s), "mg/mL" means milligram per milliliter, "L" means liter(s), "μM" means micromolar, "mM" means millimolar, "M" means molar, "mmol" means millimole(s), "μmole" means micromole(s), "kg" means kilogram(s), "g" means gram(s), "mg" means milligram(s), "μg" means microgram(s), "ng" means nanogram(s), "PCR" means polymerase chain reaction, "OD" means optical density, "OD₆₀₀" means the optical density measured at a wavelength of 600 nm, "kDa" means kilodaltons, "bp" means base pair(s), "kbp" means kilobase pair(s), "kb" means kilobase, "%" means percent, "% w/v" means weight/volume percent, "% v/v" means volume/volume percent, "HPLC" means high performance liquid chromatography, "g/L" means gram(s) per liter, "L/L" means liter(s) per liter, "ml/L" means milliliter(s) per liter, "μg/L" means microgram(s) per liter, "ng/μL" means nanogram(s) per microliter, "pmol/μL" means picomol(s) per microliter, "RPM" means rotation(s) per minute, "μmol/min/mg" means micromole(s) per minute per milligram, "mL/min" means milliliter(s) per minute, "g/L/hr" or "grams/L/hr" means grams per liter per hour, "gdcw/L" is gram dry cell weight per liter, "g/gdcw/h" is gram per gram dry cell weight per hour, "w/v" means weight per volume, "v/v" means volume per volume, "cfu/mL" means colony forming unit(s) per milliliter.

General Methods

[00238] Standard recombinant DNA and molecular cloning techniques used in the Examples are well known in the art and are described by Sambrook, et al. (Sambrook, J., Fritsch, E. F. and Maniatis, T. (Molecular Cloning: A Laboratory Manual; Cold Spring

Harbor Laboratory Press, Cold Spring Harbor, 1989) and by Ausubel, et al. (Current Protocols in Molecular Biology, Greene Publishing Assoc. and Wiley-Interscience, 1987).

[00239] Materials and methods suitable for the maintenance and growth of bacterial cultures are well known in the art. Techniques suitable for use in the following Examples may be found in Manual of Methods for General Bacteriology (Phillipp, et al., eds., American Society for Microbiology, Washington, DC., 1994) or by Thomas D. Brock (Biotechnology: A Textbook of Industrial Microbiology, Second Edition, Sinauer Associates, Inc., Sunderland, MA (1989). All reagents, restriction enzymes, and materials used for the growth and maintenance of bacterial cells were obtained from Sigma-Aldrich Chemicals (St. Louis, MO), BD Diagnostic Systems (Sparks, MD), Invitrogen (Carlsbad, CA), HiMedia (Mumbai, India), SD Fine chemicals (India), or Takara Bio Inc. (Shiga, Japan), unless otherwise specified.

[00240] The following media and stock solutions (Tables 4-7) were used in the Examples described herein.

Table 4

Yeast synthetic medium w/o amino acids and glucose (2x, base: ultrapure water)	
Component	Concentration
Yeast Nitrogen Base (YNB) w/o amino acids	13.4 g/L
Thiamine	20 mg/L
Niacin	20 mg/L
Tween & Ergosterol solution (in 50% ethanol) (10 g Ergosterol in 500 mL ethanol and 500 mL Tween® 80)	2.0 mL/L
1M MES buffer, pH=5.5	200 mL/L

[00241] Supplement amino acid solution without histidine and uracil (SAAS-1, 10x):

- 18.5 g/L synthetic complete amino acid dropout –His, -Ura (Kaiser Mixture, ForMedium™, Norfolk, United Kingdom).

[00242] Tween and Ergosterol stock solution:

- 1L Tween & Ergosterol solution contains 10 g ergosterol dissolved in 500 mL 100% ethanol and 500 mL Tween® 80 (polyoxyethylenesorbitan monooleate).

[00243] Ethanol stock solution:

- Ethanol (100%, $c(\text{C}_2\text{H}_5\text{OH}) = 17.1 \text{ M}$, 1 ml = 17.1 mmol).

- 77 -

[00244] MgCl₂ stock solution:
- 2 M MgCl₂ in bidest water.

[00245] MgSO₄ stock solution:
- 2 M MgSO₄ in bidest water.

[00246] MgCl₂ stock solution:
- 2 M CaCl₂ in bidest water.

Table 5

SEED medium	
Component	Concentration
Yeast synthetic medium w/o amino acids and with ethanol addition (2x)	50%
Supplement amino acid solution without histidine and uracil	10%
Ultrapure water	40%
Total	10 mL

Table 6

Stage 1 Medium (Base: ultrapure water)	
Component	Concentration
Yeast Nitrogen Base w/o amino acids	6.7 g/L
Yeast synthetic drop-out medium supplement without histidine and uracil	3.7 g/L
Thiamine (2 mL/L of 10 g/L stock solution)	20 mg/L
Niacin	20 mg/L
Tween & Ergosterol solution (in 50% ethanol) (10 g Ergosterol in 500 mL ethanol and 500 mL Tween® 80)	1.0 mL/L
1M MES buffer, pH=5.5	100 mL/L
Ethanol (100%)	3.5 mL/L
50% glucose (ad 3 g/L)	5.5 mL/L
Acetic acid	0.6 mL/L

Table 7

Stage 2 Medium	
Component	Concentration
Yeast Synthetic Medium w/o amino acids and glucose (2x)	50%
Amino acid solution without histidine and uracil	10%
Glucose (250 g/L)	16%
Compound stock solution (10x)	Added to each concentration (%)
Ultrapure water	to 100 %

High Performance Liquid Chromatography

[00247] Compound analysis was performed using HPLC. A Bio-Rad Aminex® HPX-87H column (Bio-Rad Laboratories, Hercules, CA) was used in an isocratic method with 0.01N sulfuric acid as eluent on an Alliance® 2695 Separations Module (Waters, Milford, MA). Flow rate was 0.60 mL/min, column temperature 40°C, injection volume 10 µL, and run time 58 min. Detection was carried out with a 2414 Refractive Index Detector (Waters, Milford, MA) operated at 40°C and an UV detector (2996 PDA; Waters, Milford, MA) at 210 nm.

Average Specific Consumption and Production Rate(s)

[00248] Average specific consumption and production rate(s) [q(ave)] were calculated by determining the concentration change of a substrate (s) or a product (p) during a time interval and dividing it by the average biomass concentration during this time interval. During exponential growth or biomass decrease at the specific growth rate (μ), the average biomass concentration [cx(ave)] in a time interval starting at time point t_1 and ending at time point t_2 was determined according to $cx(ave) = (cx(t_2) - cx(t_1))/(t_2 - t_1)/\mu$. In all other situations, the average biomass concentration cx(ave) was determined according to $cx(ave) = (cx(t_1) + cx(t_2))/2$.

EXAMPLE 1**Construction of a *Saccharomyces cerevisiae* Strain PNY 2068**

[00249] *Saccharomyces cerevisiae* strain PNY0827 is used as the host cell for further genetic manipulation. PNY0827 refers to a strain derived from *Saccharomyces cerevisiae* which has been deposited at the ATCC under the Budapest Treaty on September 22, 2011 at the American Type Culture Collection, Patent Depository 10801 University Boulevard, Manassas, VA 20110-2209 and has the patent deposit designation PTA-12105.

Deletion of URA3 and sporulation into haploids

[00250] In order to delete the endogenous URA3 coding region, a deletion cassette was PCR-amplified from pLA54 (SEQ ID NO: 129) which contains a P_{TEF1} -*kanMX4*-*TEF1*t cassette flanked by loxP sites to allow homologous recombination in vivo and subsequent removal of the KANMX4 marker. PCR was performed using Phusion® High Fidelity PCR Master Mix (New England BioLabs, Ipswich, MA) and primers BK505 (SEQ ID NO: 130) and BK506 (SEQ ID NO: 131). The *URA3* portion of each primer was derived from the 5' region 180 bp upstream of the *URA3* ATG and 3' region 78 bp downstream of the coding region such that integration of the *kanMX4* cassette results in replacement of the *URA3* coding region. The PCR product was transformed into PNY0827 using standard genetic techniques (Methods in Yeast Genetics, 2005, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 201-202) and transformants were selected on YEP medium supplemented 2% glucose and 100 µg/ml Geneticin at 30°C. Transformants were screened by colony PCR with primers LA468 (SEQ ID NO: 132) and LA492 (SEQ ID NO: 133) to verify presence of the integration cassette. A heterozygous diploid was obtained: NYLA98, which has the genotype MATa/ α URA3/*ura3::loxP-kanMX4-loxP*. To obtain haploids, NYLA98 was sporulated using standard methods (Codón, et al., Appl. Environ. Microbiol. 61:630, 1995). Tetrads were dissected using a micromanipulator and grown on rich YPE medium supplemented with 2% glucose. Tetrads containing four viable spores were patched onto synthetic complete medium lacking uracil supplemented with 2% glucose, and the mating type was verified by multiplex colony PCR using primers AK109-1 (SEQ ID NO: 134), AK109-2 (SEQ ID NO: 135), and AK109-3 (SEQ ID NO: 136). The resulting haploid

strain called NYLA103, which has the genotype: MAT α *ura3* Δ ::loxP-*kanMX4*-loxP, and NYLA106, which has the genotype: MAT α *ura3* Δ ::loxP-*kanMX4*-loxP.

Deletion of His3

[00251] To delete the endogenous *HIS3* coding region, a scarless deletion cassette was used. The four fragments for the PCR cassette for the scarless *HIS3* deletion were amplified using Phusion® High Fidelity PCR Master Mix (New England BioLabs, Ipswich, MA) and CEN.PK 113-7D genomic DNA as template, prepared with a Gentra® Puregene® Yeast/Bact kit (Qiagen, Valencia, CA). *HIS3* Fragment A was amplified with primer oBP452 (SEQ ID NO: 137) and primer oBP453 (SEQ ID NO: 138), containing a 5' tail with homology to the 5' end of *HIS3* Fragment B. *HIS3* Fragment B was amplified with primer oBP454 (SEQ ID NO: 139), containing a 5' tail with homology to the 3' end of *HIS3* Fragment A, and primer oBP455 (SEQ ID NO: 140) containing a 5' tail with homology to the 5' end of *HIS3* Fragment U. *HIS3* Fragment U was amplified with primer oBP456 (SEQ ID NO: 141), containing a 5' tail with homology to the 3' end of *HIS3* Fragment B, and primer oBP457 (SEQ ID NO: 142), containing a 5' tail with homology to the 5' end of *HIS3* Fragment C. *HIS3* Fragment C was amplified with primer oBP458 (SEQ ID NO: 143), containing a 5' tail with homology to the 3' end of *HIS3* Fragment U, and primer oBP459 (SEQ ID NO: 144). PCR products were purified with a PCR purification kit (Qiagen, Valencia, CA). *HIS3* Fragment AB was created by overlapping PCR by mixing *HIS3* Fragment A and *HIS3* Fragment B and amplifying with primers oBP452 (SEQ ID NO: 137) and oBP455 (SEQ ID NO: 140). *HIS3* Fragment UC was created by overlapping PCR by mixing *HIS3* Fragment U and *HIS3* Fragment C and amplifying with primers oBP456 (SEQ ID NO: 141) and oBP459 (SEQ ID NO: 144). The resulting PCR products were purified on an agarose gel followed by a gel extraction kit (Qiagen, Valencia, CA). The *HIS3* ABUC cassette was created by overlapping PCR by mixing *HIS3* Fragment AB and *HIS3* Fragment UC and amplifying with primers oBP452 (SEQ ID NO: 137) and oBP459 (SEQ ID NO: 144). The PCR product was purified with a PCR purification kit (Qiagen, Valencia, CA). Competent cells of NYLA106 were transformed with the *HIS3* ABUC PCR cassette and were plated on synthetic complete medium lacking uracil supplemented with 2% glucose at 30°C. Transformants were screened to verify correct integration by replica plating onto synthetic complete medium lacking histidine and supplemented with 2% glucose at 30°C.

Genomic DNA preps were made to verify the integration by PCR using primers oBP460 (SEQ ID NO: 145) and LA135 (SEQ ID NO: 146) for the 5' end and primers oBP461 (SEQ ID NO: 147) and LA92 (SEQ ID NO: 148) for the 3' end. The *URA3* marker was recycled by plating on synthetic complete medium supplemented with 2% glucose and 5-FOA at 30°C following standard protocols. Marker removal was confirmed by patching colonies from the 5-FOA plates onto SD -URA medium to verify the absence of growth. The resulting identified strain, called PNY2003 has the genotype: MATa *ura3Δ::loxP-kanMX4-loxP his3Δ*.

Deletion of PDC1

[00252] To delete the endogenous *PDC1* coding region, a deletion cassette was PCR-amplified from pLA59 (SEQ ID NO: 149), which contains a *URA3* marker flanked by degenerate loxP sites to allow homologous recombination in vivo and subsequent removal of the *URA3* marker. PCR was done by using Phusion® High Fidelity PCR Master Mix (New England BioLabs, Ipswich, MA) and primers LA678 (SEQ ID NO: 150) and LA679 (SEQ ID NO: 151). The *PDC1* portion of each primer was derived from the 5' region 50 bp downstream of the *PDC1* start codon and 3' region 50 bp upstream of the stop codon such that integration of the *URA3* cassette results in replacement of the *PDC1* coding region but leaves the first 50 bp and the last 50 bp of the coding region. The PCR product was transformed into PNY2003 using standard genetic techniques and transformants were selected on synthetic complete medium lacking uracil and supplemented with 2% glucose at 30°C. Transformants were screened to verify correct integration by colony PCR using primers LA337 (SEQ ID NO: 152), external to the 5' coding region and LA135 (SEQ ID NO: 146), an internal primer to *URA3*. Positive transformants were then screened by colony PCR using primers LA692 (SEQ ID NO: 153) and LA693 (SEQ ID NO: 154), internal to the *PDC1* coding region. The *URA3* marker was recycled by transforming with pLA34 (SEQ ID NO: 155) containing the CRE recombinase under the *GALI* promoter and plated on synthetic complete medium lacking histidine and supplemented with 2% glucose at 30°C. Transformants were plated on rich medium supplemented with 0.5% galactose to induce the recombinase. Marker removal was confirmed by patching colonies to synthetic complete medium lacking uracil and supplemented with 2% glucose to verify absence of growth. The

- 82 -

resulting identified strain, called PNY2008 has the genotype: MATa *ura3Δ::loxP-kanMX4-loxP his3Δ pdc1Δ::loxP71/66*.

Deletion of PDC5

[00253] To delete the endogenous *PDC5* coding region, a deletion cassette was PCR-amplified from pLA59 (SEQ ID NO: 149), which contains a *URA3* marker flanked by degenerate loxP sites to allow homologous recombination in vivo and subsequent removal of the *URA3* marker. PCR was done by using Phusion® High Fidelity PCR Master Mix (New England BioLabs, Ipswich, MA) and primers LA722 (SEQ ID NO: 156) and LA733 (SEQ ID NO: 157). The *PDC5* portion of each primer was derived from the 5' region 50 bp upstream of the *PDC5* start codon and 3' region 50 bp downstream of the stop codon such that integration of the *URA3* cassette results in replacement of the entire *PDC5* coding region. The PCR product was transformed into PNY2008 using standard genetic techniques and transformants were selected on synthetic complete medium lacking uracil and supplemented with 1% ethanol at 30°C. Transformants were screened to verify correct integration by colony PCR using primers LA453 (SEQ ID NO: 158), external to the 5' coding region and LA135 (SEQ ID NO: 146), an internal primer to *URA3*. Positive transformants were then screened by colony PCR using primers LA694 (SEQ ID NO: 159) and LA695 (SEQ ID NO: 160), internal to the *PDC5* coding region. The *URA3* marker was recycled by transforming with pLA34 (SEQ ID NO: 155) containing the CRE recombinase under the *GAL1* promoter and plated on synthetic complete medium lacking histidine and supplemented with 1% ethanol at 30°C. Transformants were plated on rich YEP medium supplemented with 1% ethanol and 0.5% galactose to induce the recombinase. Marker removal was confirmed by patching colonies to synthetic complete medium lacking uracil and supplemented with 1% ethanol to verify absence of growth. The resulting identified strain, called PNY2009 has the genotype: MATa *ura3Δ::loxP-kanMX4-loxP his3Δ pdc1Δ::loxP71/66 pdc5Δ::loxP71/66*.

Deletion of FRA2

[00254] The *FRA2* deletion was designed to delete 250 nucleotides from the 3' end of the coding sequence, leaving the first 113 nucleotides of the *FRA2* coding sequence intact. An in-frame stop codon was present seven nucleotides downstream of the deletion. The four

fragments for the PCR cassette for the scarless *FRA2* deletion were amplified using Phusion® High Fidelity PCR Master Mix (New England BioLabs, Ipswich, MA) and CEN.PK 113-7D genomic DNA as template, prepared with a Gentra® Puregene® Yeast/Bact kit (Qiagen, Valencia, CA). *FRA2* Fragment A was amplified with primer oBP594 (SEQ ID NO: 161) and primer oBP595 (SEQ ID NO: 162), containing a 5' tail with homology to the 5' end of *FRA2* Fragment B. *FRA2* Fragment B was amplified with primer oBP596 (SEQ ID NO: 163), containing a 5' tail with homology to the 3' end of *FRA2* Fragment A, and primer oBP597 (SEQ ID NO: 164), containing a 5' tail with homology to the 5' end of *FRA2* Fragment U. *FRA2* Fragment U was amplified with primer oBP598 (SEQ ID NO: 165), containing a 5' tail with homology to the 3' end of *FRA2* Fragment B, and primer oBP599 (SEQ ID NO: 166), containing a 5' tail with homology to the 5' end of *FRA2* Fragment C. *FRA2* Fragment C was amplified with primer oBP600 (SEQ ID NO: 167), containing a 5' tail with homology to the 3' end of *FRA2* Fragment U, and primer oBP601 (SEQ ID NO: 168). PCR products were purified with a PCR purification kit (Qiagen, Valencia, CA). *FRA2* Fragment AB was created by overlapping PCR by mixing *FRA2* Fragment A and *FRA2* Fragment B and amplifying with primers oBP594 (SEQ ID NO: 161) and oBP597 (SEQ ID NO: 164). *FRA2* Fragment UC was created by overlapping PCR by mixing *FRA2* Fragment U and *FRA2* Fragment C and amplifying with primers oBP598 (SEQ ID NO: 165) and oBP601 (SEQ ID NO: 168). The resulting PCR products were purified on an agarose gel followed by a gel extraction kit (Qiagen, Valencia, CA). The *FRA2* ABUC cassette was created by overlapping PCR by mixing *FRA2* Fragment AB and *FRA2* Fragment UC and amplifying with primers oBP594 (SEQ ID NO: 161) and oBP601 (SEQ ID NO: 168). The PCR product was purified with a PCR purification kit (Qiagen, Valencia, CA).

[00255] To delete the endogenous *FRA2* coding region, the scarless deletion cassette obtained above was transformed into PNY2009 using standard techniques and plated on synthetic complete medium lacking uracil and supplemented with 1% ethanol. Genomic DNA preps were made to verify the integration by PCR using primers oBP602 (SEQ ID NO: 169) and LA135 (SEQ ID NO: 146) for the 5' end, and primers oBP602 (SEQ ID NO: 169) and oBP603 (SEQ ID NO: 170) to amplify the whole locus. The *URA3* marker was recycled by plating on synthetic complete medium supplemented with 1% ethanol and 5-FOA (5-Fluoroorotic Acid) at 30°C following standard protocols. Marker removal was confirmed by

- 84 -

patching colonies from the 5-FOA plates onto synthetic complete medium lacking uracil and supplemented with 1% ethanol to verify the absence of growth. The resulting identified strain, PNY2037, has the genotype: MATa *ura3Δ::loxP-kanMX4-loxP his3Δ pdc1Δ::loxP71/66 pdc5Δ::loxP71/66 fra2Δ*.

Addition of native 2 micron plasmid

[00256] The loxP71-URA3-loxP66 marker was PCR-amplified using Phusion® DNA polymerase (New England BioLabs, Ipswich, MA) from pLA59 (SEQ ID NO: 149), and transformed along with the LA811x817 (SEQ ID NOs: 171, 172) and LA812x818 (SEQ ID NOs: 173, 174) 2-micron plasmid fragments into strain PNY2037 on SE –URA plates at 30°C. The resulting strain PNY2037 2μ::loxP71-URA3-loxP66 was transformed with pLA34 (pRS423::cre) (SEQ ID NO: 155) and selected on SE –HIS –URA plates at 30°C. Transformants were patched onto YP-1% galactose plates and allowed to grow for 48 hr at 30°C to induce Cre recombinase expression. Individual colonies were then patched onto SE –URA, SE –HIS, and YPE plates to confirm URA3 marker removal. The resulting identified strain, PNY2050, has the genotype: MATa *ura3Δ::loxP-kanMX4-loxP, his3Δ pdc1Δ::loxP71/66 pdc5Δ::loxP71/66 fra2Δ* 2-micron.

Deletion of GPD2

[00257] To delete the endogenous *GPD2* coding region, a deletion cassette was PCR-amplified from pLA59 (SEQ ID NO: 149), which contains a *URA3* marker flanked by degenerate loxP sites to allow homologous recombination in vivo and subsequent removal of the *URA3* marker. PCR was done by using Phusion® High Fidelity PCR Master Mix (New England BioLabs, Ipswich, MA) and primers LA512 (SEQ ID NO: 175) and LA513 (SEQ ID NO: 176). The *GPD2* portion of each primer was derived from the 5' region 50 bp upstream of the *GPD2* start codon and 3' region 50 bp downstream of the stop codon such that integration of the *URA3* cassette results in replacement of the entire *GPD2* coding region. The PCR product was transformed into PNY2050 using standard genetic techniques and transformants were selected on synthetic complete medium lacking uracil and supplemented with 1% ethanol at 30°C. Transformants were screened to verify correct integration by colony PCR using primers LA516 (SEQ ID NO: 177), external to the 5' coding

region and LA135 (SEQ ID NO: 146), internal to *URA3*. Positive transformants were then screened by colony PCR using primers LA514 (SEQ ID NO: 178) and LA515 (SEQ ID NO: 179), internal to the *GPD2* coding region. The *URA3* marker was recycled by transforming with pLA34 (SEQ ID NO: 155) containing the CRE recombinase under the *GAL1* promoter and plated on synthetic complete medium lacking histidine and supplemented with 1% ethanol at 30°C. Transformants were plated on rich medium supplemented with 1% ethanol and 0.5% galactose to induce the recombinase. Marker removal was confirmed by patching colonies to synthetic complete medium lacking uracil and supplemented with 1% ethanol to verify absence of growth. The resulting identified strain, PNY2056, has the genotype: MATa *ura3Δ::loxP-kanMX4-loxP his3Δ pdc1Δ::loxP71/66 pdc5Δ::loxP71/66 fra2Δ 2-micron gpd2Δ*.

Deletion of YMR226 and integration of AlsS

[00258] To delete the endogenous *YMR226C* coding region, an integration cassette was PCR-amplified from pLA71 (SEQ ID NO: 180), which contains the gene acetolactate synthase from the species *Bacillus subtilis* with a *FBA1* promoter and a *CYC1* terminator, and a *URA3* marker flanked by degenerate loxP sites to allow homologous recombination in vivo and subsequent removal of the *URA3* marker. PCR was done by using KAPA HiFi™ (Kapa Biosystems, Woburn, MA) and primers LA829 (SEQ ID NO: 181) and LA834 (SEQ ID NO: 182). The *YMR226C* portion of each primer was derived from the first 60 bp of the coding sequence and 65 bp that are 409 bp upstream of the stop codon. The PCR product was transformed into PNY2056 using standard genetic techniques and transformants were selected on synthetic complete medium lacking uracil and supplemented with 1% ethanol at 30°C. Transformants were screened to verify correct integration by colony PCR using primers N1257 (SEQ ID NO: 183), external to the 5' coding region and LA740 (SEQ ID NO: 184), internal to the *FBA1* promoter. Positive transformants were then screened by colony PCR using primers N1257 (SEQ ID NO: 183) and LA830 (SEQ ID NO: 185), internal to the *YMR226C* coding region, and primers LA830 (SEQ ID NO: 185), external to the 3' coding region, and LA92 (SEQ ID NO: 148), internal to the *URA3* marker. The *URA3* marker was recycled by transforming with pLA34 (SEQ ID NO: 155) containing the CRE recombinase under the *GAL1* promoter and plated on synthetic complete medium lacking histidine and supplemented with 1% ethanol at 30°C. Transformants were plated on rich medium

- 86 -

supplemented with 1% ethanol and 0.5% galactose to induce the recombinase. Marker removal was confirmed by patching colonies to synthetic complete medium lacking uracil and supplemented with 1% ethanol to verify absence of growth. The resulting identified strain, PNY2061, has the genotype: MATa *ura3Δ::loxP-kanMX4-loxP his3Δ pdc1Δ::loxP71/66 pdc5Δ::loxP71/66 fra2Δ 2-micron gpd2Δ ymr226cΔ::P_{FBA1}-alsS_Bs-CYC1t-loxP71/66*.

Deletion of ALD6 and integration of KivD

[00259] To delete the endogenous *ALD6* coding region, an integration cassette was PCR-amplified from pLA78 (SEQ ID NO: 186), which contains the *kivD* gene from the species *Listeria grayi* with a hybrid *FBA1* promoter and a *TDH3* terminator, and a *URA3* marker flanked by degenerate loxP sites to allow homologous recombination in vivo and subsequent removal of the *URA3* marker. PCR was done by using KAPA HiFi™ (Kapa Biosystems, Woburn, MA) and primers LA850 (SEQ ID NO: 187) and LA851 (SEQ ID NO: 188). The *ALD6* portion of each primer was derived from the first 65 bp of the coding sequence and the last 63 bp of the coding region. The PCR product was transformed into PNY2061 using standard genetic techniques and transformants were selected on synthetic complete medium lacking uracil and supplemented with 1% ethanol at 30°C. Transformants were screened to verify correct integration by colony PCR using primers N1262 (SEQ ID NO: 189), external to the 5' coding region and LA740 (SEQ ID NO: 184), internal to the *FBA1* promoter. Positive transformants were then screened by colony PCR using primers N1263 (SEQ ID NO: 190), external to the 3' coding region, and LA92 (SEQ ID NO: 148), internal to the *URA3* marker. The *URA3* marker was recycled by transforming with pLA34 (SEQ ID NO: 155) containing the CRE recombinase under the *GALI* promoter and plated on synthetic complete medium lacking histidine and supplemented with 1% ethanol at 30°C. Transformants were plated on rich medium supplemented with 1% ethanol and 0.5% galactose to induce the recombinase. Marker removal was confirmed by patching colonies to synthetic complete medium lacking uracil and supplemented with 1% ethanol to verify absence of growth. The resulting identified strain, PNY2065, has the genotype: MATa *ura3Δ::loxP-kanMX4-loxP his3Δ pdc1Δ::loxP71/66 pdc5Δ::loxP71/66 fra2Δ 2-micron gpd2Δ ymr226cΔ::P_{FBA1}-alsS_Bs-CYC1t-loxP71/66 ald6Δ::(UAS)PGK1-P_{FBA1}-kivD_Lg-TDH3t-loxP71*.

Deletion of ADH1 and integration of ADH

[00260] ADH1 is the endogenous alcohol dehydrogenase present in *Saccharomyces cerevisiae*. As described below, the endogenous ADH1 was replaced with alcohol dehydrogenase (ADH) from *Beijerinckii indica*.

[00261] To delete the endogenous ADH1 coding region, an integration cassette was PCR-amplified from pLA65 (SEQ ID NO: 191), which contains the alcohol dehydrogenase from the species *Beijerinckii indica* with an *ILV5* promoter and a *ADH1* terminator, and a *URA3* marker flanked by degenerate loxP sites to allow homologous recombination in vivo and subsequent removal of the *URA3* marker. PCR was done by using KAPA HiF™ (Kapa Biosystems, Woburn, MA) and primers LA855 (SEQ ID NO: 192) and LA856 (SEQ ID NO: 193). The *ADH1* portion of each primer was derived from the 5' region 50 bp upstream of the *ADH1* start codon and the last 50 bp of the coding region. The PCR product was transformed into PNY2065 using standard genetic techniques and transformants were selected on synthetic complete medium lacking uracil and supplemented with 1% ethanol at 30°C. Transformants were screened to verify correct integration by colony PCR using primers LA414 (SEQ ID NO: 194), external to the 5' coding region and LA749 (SEQ ID NO: 195), internal to the *ILV5* promoter. Positive transformants were then screened by colony PCR using primers LA413 (SEQ ID NO: 196), external to the 3' coding region, and LA92 (SEQ ID NO: 148), internal to the *URA3* marker. The *URA3* marker was recycled by transforming with pLA34 (SEQ ID NO: 155) containing the CRE recombinase under the *GAL1* promoter and plated on synthetic complete medium lacking histidine and supplemented with 1% ethanol at 30°C. Transformants were plated on rich medium supplemented with 1% ethanol and 0.5% galactose to induce the recombinase. Marker removal was confirmed by patching colonies to synthetic complete medium lacking uracil and supplemented with 1% ethanol to verify absence of growth. The resulting identified strain, called PNY2066 has the genotype: MATa *ura3Δ::loxP-kanMX4-loxP his3Δ pdc1Δ::loxP71/66 pdc5Δ::loxP71/66 fra2Δ 2-micron gpd2Δ ymr226cΔ::P_{FBA1}-alsS_Bs-CYC1t-loxP71/66 ald6Δ::(UAS)PGK1-P_{FBA1}-kivD_Lg-TDH3t-loxP71/66 adh1Δ::P_{ILV5}-ADH_Bi(y)-ADH1t-loxP71/66*.

Integration of ADH into *pdcl*Δ locus

[00262] To integrate an additional copy of *ADH* at the *pdcl*Δ region, an integration cassette was PCR-amplified from pLA65 (SEQ ID NO: 192), which contains the alcohol dehydrogenase from the species *Beijerinckii indica* with an *ADH1* terminator, and a *URA3* marker flanked by degenerate loxP sites to allow homologous recombination in vivo and subsequent removal of the *URA3* marker. PCR was done by using KAPA HiFi™ (Kapa Biosystems, Woburn, MA) and primers LA860 (SEQ ID NO: 197) and LA679 (SEQ ID NO: 151). The *PDC1* portion of each primer was derived from the 5' region 60 bp upstream of the *PDC1* start codon and 50 bp that are 103 bp upstream of the stop codon. The endogenous *PDC1* promoter was used. The PCR product was transformed into PNY2066 using standard genetic techniques and transformants were selected on synthetic complete medium lacking uracil and supplemented with 1% ethanol at 30°C. Transformants were screened to verify correct integration by colony PCR using primers LA337 (SEQ ID NO: 152), external to the 5' coding region and N1093 (SEQ ID NO: 198), internal to the BiADH gene. Positive transformants were then screened by colony PCR using primers LA681 (SEQ ID NO: 199), external to the 3' coding region, and LA92 (SEQ ID NO: 148), internal to the *URA3* marker. The *URA3* marker was recycled by transforming with pLA34 (SEQ ID NO: 155) containing the CRE recombinase under the *GAL1* promoter and plated on synthetic complete medium lacking histidine and supplemented with 1% ethanol at 30°C. Transformants were plated on rich medium supplemented with 1% ethanol and 0.5% galactose to induce the recombinase. Marker removal was confirmed by patching colonies to synthetic complete medium lacking uracil and supplemented with 1% ethanol to verify absence of growth. The resulting identified strain, called PNY2068 has the genotype: MATa *ura3*Δ::loxP-kanMX4-loxP *his3*Δ *pdcl*Δ::loxP71/66 *pdcl*Δ::loxP71/66 *fra2*Δ 2-micron *gpd2*Δ *ymr226c*Δ::P_{FBA1}-*alsS*_{Bs}-*CYC1*t-loxP71/66 *ald6*Δ::(UAS)*PGK1*-P_{FBA1}-*kivD*_{Lg}-*TDH3*t-loxP71/66 *adh1*Δ::P_{ILV5}-*ADH*_{Bi(y)}-*ADH1*t-loxP71/66 *pdcl*Δ::P_{PDC1}-*ADH*_{Bi(y)}-*ADH1*t-loxP71/66.

EXAMPLE 2**Construction of a *Saccharomyces cerevisiae* Strain PNY2071**

[00263] Strain PNY2071 has the genomic background MATa *ura3Δ::loxP* *his3Δ pdc5Δ::loxP66/71* *fra2Δ* 2-micron plasmid (CEN.PK2) *gpd2Δ::loxP71/66* *ymr226CA::P[FBA1]-ALS|alsS_Bs-CYC1t-loxP71/66* *ald6Δ::UAS(PGK1)P[FBA1]-KIVD|Lg(y)-TDH3t-loxP71/66* *adh1Δ::P[ILV5]-ADH|Bi(y)-ADHt-loxP71/66* *pdc1Δ::P[PDC1]-ADH|Bi(y)-ADHt-loxP71/66*.

[00264] PNY2071 was generated by transforming PNY2068 with plasmids pHR81-K9D3 and pYZ067DkivDDadh. Plasmid pHR81-K9D3 (SEQ ID NO. 200) and plasmid pYZ067DkivDDadh (SEQ ID NO. 201) are described in, for example, U.S. Patent Application Publication No. 2012/0208246, the entire contents of which are herein incorporated by reference.

EXAMPLE 3**Effects of Magnesium Supplementation on Isobutanol Production**

[00265] A 125 mL aerobic shake flask was prepared with 10 mL SEED medium (Table 5) and inoculated with a vial of frozen glycerol stock culture of PNY2071. The culture was incubated at 30°C and 250 rpm for 24 h in an Innova Laboratory Shaker (New Brunswick Scientific, Edison, NJ). The seed culture (5 mL) was transferred to 500 mL aerobic shake flasks filled with 95 mL STAGE 1 medium (Table 6) to give a total culture volume of 100 mL and incubated again at 250 rpm for 24 h. Sufficient culture volume to yield an initial OD of approximately 1.0 was transferred to 50 mL sterile centrifuge tubes, centrifuged at 9500 rpm for 20 min. The supernatants were discarded and the cell pellets re-suspended in appropriate volumes of STAGE 2 medium (Table 7) with amino acids. Respective amounts of MgCl₂ stock solution and bidest water were added to give a total volume of 12 mL. The cell cultures (12 mL) were transferred to each 25 ml Balch tube. Each Balch tube was fitted with a butyl rubber septum and crimped to the tube with a sheet metal with circular opening to allow samples withdrawal by syringes. Growth of the cell was monitored by OD measurements. Optical density was measured with an Ultrospec™ 3000 spectrophotometer (Pharmacia Biotech/GE Healthcare Biosciences, Pittsburgh, PA) at $\lambda = 600$ nm. Cell dry

weight concentration was calculated from the OD readings assuming an OD-DW-correlation of 0.33 gDW/OD. Balch tube experiments were conducted for 48 h.

[00266] Extracellular compound analysis in supernatant was accomplished by HPLC. An Aminex® HPX-87H column (Bio-Rad, Hercules, CA) was used in an isocratic method with 0.01N sulfuric acid as eluent on an Alliance® 2695 Separations Module (Waters Corp., Milford, MA). Flow rate was 0.60 mL/min, column temperature 40°C, injection volume 10 µL and run time 58 min. Detection was carried out with a refractive index detector (Waters 2414 RI, Waters Corp., Milford, MA) operated at 40°C and an UV detector (Waters 2996 PDA, Waters Corp., Milford, MA) at 210 nm.

[00267] Specific maximum growth rates of PNY2071 cultures were determined during aerobic growth in YNB-based synthetic medium with and without additional supplementation of either 0.2 and 0.4 M MgCl₂. Supplementation of MgCl₂ resulted in an increased specific isobutanol production rate as compared to the non-supplemented cultures. Results are shown in Figure 1.

[00268] Specific maximum growth rates and isobutanol titers of PNY2071 cultures were determined during aerobic growth in YNB-based synthetic medium with and without additional supplementation of MgCl₂ in concentrations of 0.05 M (50 mM) to 0.30 M (300 mM). PNY2071 cultures were grown as described herein. Cultures supplemented with magnesium exhibited increased biomass production compared to non-supplemented cultures. Results are shown in Figure 2.

[00269] Final isobutanol titers in supplemented cultures were higher as compared to non-supplemented cultures. Results are shown in Figure 3. The higher final isobutanol titers in the supplemented cultures were not only an effect of the improved growth of the cultures, but also due to higher specific isobutanol production rates as shown in Figure 4. Supplementing cultures with magnesium in the range 0.05 to 0.25 M resulted in increased final isobutanol titers. The elevated final isobutanol titers resulted from a combination of factors such as improved biomass formation, higher specific isobutanol production rates, and higher product yields.

[00270] To validate the positive effect from magnesium supplementation, MgCl₂ or MgSO₄ were added to the cultures to yield similar concentrations of Mg²⁺. Final isobutanol titers of cultures supplemented with either MgCl₂ or with MgSO₄ demonstrated similar results as shown in Figure 5.

[00271] Final isobutanol titers in cultures supplemented with magnesium and calcium indicated that high ratios of calcium-to-magnesium may interfere with isobutanol production. Results are shown in Figure 6. It may be beneficial to maintain lower calcium-to-magnesium ratios in isobutanol-producing cultures, for example, by removing calcium from the medium by precipitation or ion exchange chromatography or by supplementing the medium with magnesium.

EXAMPLE 4

Effects of Magnesium Supplementation on Isobutanol and Byproduct Production

[00272] Isobutanol and byproduct yields of PNY2071 cultures were determined during growth in YNB-based synthetic medium with and without additional supplementation of MgCl_2 in concentrations of 0.05 M (50 μM) to 0.30 M (300 μM). PNY2071 cultures were grown as described in Example 3. Growth measurements and extracellular compound analysis were conducted as described in Example 3.

[00273] Analysis of isobutanol yield and byproduct spectrum showed increased isobutanol and increased glycerol formation in cultures supplemented with magnesium compared to non-supplemented cultures (data not shown). The yield increase in the supplemented cultures may be partly explained by decreased formation of 2,3-dihydroxyisovalerate (DHIV) as shown in Figure 7. A concentration time profile for isobutanol and DHIV concentration in cultures with and without magnesium supplementation demonstrated that the positive effects of magnesium supplementation are observed throughout growth (or production) phase. Results are as shown in Figure 8. The enzyme dihydroxyacid dehydratase (DHAD) catalyzes the conversion of 2,3-DHIV to α -ketoisovalerate. The results shown in Figure 8 suggest that DHAD activity is increased in cultures supplemented with magnesium.

EXAMPLE 5

Effects of Magnesium Supplementation on Mash

[00274] A 125 mL aerobic shake flask was prepared with 10 mL SEED medium (Table 5) and inoculated with a vial of frozen glycerol stock culture of PNY2071. The culture was incubated at 30°C and 250 rpm for 24 h in an Innova Laboratory Shaker (New Brunswick Scientific, Edison, NJ). The seed culture (5 mL) was transferred to 500 mL aerobic shake

- 92 -

flasks filled with 95 mL STAGE 1 medium (Table 6) to give a total culture volume of 100 mL and incubated again at 250 rpm for 24 h. Sufficient culture volume to yield an initial OD of approximately 1.0 was transferred to 50 mL sterile centrifuge tubes, and centrifuged at 9500 rpm for 20 min. The supernatants were discarded and the cell pellets re-suspended in appropriate volumes of corn mash medium (Table 8). Respective amounts of test solutions were added to give a total volume of 12 mL. The cell cultures (12 mL) were transferred to each 25 ml Balch tube. Each Balch tube was fitted with a butyl rubber septum and crimped to the tube with a sheet metal with circular opening to allow samples withdrawal by syringes. Performance of the cultures were monitored by measuring substrate and product concentration using HPLC and glucose concentrations were measured by HPLC and enzyme assay.

Table 8

Corn Mash Medium	
Component	Concentration
Centrifuged corn mash	168.30 mL
Urea stock solution	0.80 mL
Nicotinic acid (10 g/L) + thiamine (10 g/L) solution	0.60 mL
Ethanol	0.12 mL
Glucose Solution	10 mL
Ergosterol & Tween solution	0.20 mL
1 M MES buffer (pH = 5.5)	20 mL

[00275] Compound analysis in supernatant was accomplished by HPLC. An Aminex® HPX-87H column (Bio-Rad, Hercules, CA) was used in an isocratic method with 0.01N sulfuric acid as eluent on an Alliance® 2695 Separations Module (Waters Corp., Milford, MA). Flow rate was 0.60 mL/min, column temperature 40°C, injection volume 10 µL and run time 58 min. Detection was carried out with a refractive index detector (Waters 2414 RI, Waters Corp., Milford, MA) operated at 40°C and an UV detector (Waters 2996 PDA, Waters Corp., Milford, MA) at 210 nm.

[00276] Corn mash medium was supplemented with magnesium and glucose. Final isobutanol titers in supplemented cultures were higher as compared to non-supplemented

cultures. Results are shown in Figure 9. Comparing the isobutanol production of a non-supplemented culture with a culture supplemented with 0.05 M MgCl_2 , significant differences in performance were observed between the supplemented and non-supplemented cultures. Results are shown in Figure 10. An increase in glycerol formation was also observed in the supplemented cultures (data not shown). During the time course of fermentation, a continuous increase in the ratio of isobutanol produced as compared to glycerol.

EXAMPLE 6

Supplementation with Backset

[00277] A *Saccharomyces cerevisiae* strain that was engineered to produce isobutanol (isobutanologen) or a *Saccharomyces cerevisiae* strain that produces ethanol from a carbohydrate source (ethanologen), was grown in defined medium (Difco™ Yeast Nitrogen Base without amino acids 6.7 g/L, Ref No. 291920; ForMedium™ Synthetic Complete Drop-out (Kaiser Mixture, Norfolk, United Kingdom) -His, -Ura 3.7 g/L, Ref No. DSCK10015; MES Buffer 19.5 g/L, P/N M3671); dextrose 30 g/L). The pH of the medium was adjusted to 5.8-6.2 using sodium hydroxide. The cultures were started in a seed flask (500 mL defined medium in a 2 L, baffled, vented shake flask) by adding a portion of a thawed vial to the flask at 29-31°C in an incubator rotating at 260-300 rpm and grown to a final biomass concentration of $1\text{--}2 \times 10^7$ cfu/mL (isobutanologen) or $10\text{--}30 \times 10^7$ cfu/mL (ethanologen).

Liquefied Mash Preparation without Backset

[00278] The components (27-33 wt% wet corn ground through a 1 mm screen, 67-73 wt% tap water, and alpha-amylase) for making liquefied mash were added to a pot at 20-55°C, mixed with a mechanical stirrer, heated to 85°C, held for 60-120 min, and then cooled to < 59°C. The material was transferred to centrifuge bottles, centrifuged in a Sorval® centrifuge (RC-5B, RC-5C, RC-3C) for 45 min at 5000-8000 rpm using a 4 x 1L or 6 x 500 mL fixed angle rotor. All material (thin mash) except for the wet pellet was transferred to 1 L bottles at 600-800 mL per bottle. Each bottle of thin mash was autoclaved for a 30 min, 121°C liquid sterilization cycle with the caps loosened. The bottles were removed

from the autoclave after the cycle and allowed to cool in a sterile bio-hood. The bottle caps are then sealed and the material was stored at in a refrigerator until needed.

Liquefied Mash Preparation with Backset

[00279] The components for making liquefied mash were: 27-33 wt% wet corn ground through a 1 mm screen, 67-73 wt% tap water, backset, (50-99 water volume % tap water and 1-50 water volume % thin stillage (backset) from a commercial-scale ethanol plant), and alpha-amylase. These components were added to a pot at 20-55°C, mixed with a mechanical stirrer, heated to 85°C, held for 60-120 min, and then cooled to < 59°C. The material was transferred to centrifuge bottles, centrifuged in a Sorval® centrifuge (RC-5B, RC-5C, RC-3C) for 45 min at 5000-8000 rpm using a 4 x 1L or 6 x 500 mL fixed angle rotor. All material except for the wet pellet (thin mash) was transferred to 1 L bottles at 600-800 mL per bottle. Each bottle of thin mash was autoclaved for 30 min, 121°C liquid sterilization cycle with the caps loosened. The bottles were removed from the autoclave after the cycle and allowed to cool in a sterile bio-hood. The bottle caps were then sealed and the material was stored in a refrigerator until needed.

Initial Fermentation Vessel Preparation

[00280] A 3 L fermentation vessel (Sartorius AG, Goettingen, Germany BioStat B+ Control unit with an applikon® Biotechnology glass vessel, Dover, NJ) was charged with medium (e.g., liquefied mash with or without backset). A pH probe was calibrated through the Sartorius controller. The zero was calibrated at pH=7. The span was calibrated at pH=4. The probe was then placed into the fermentation vessel. In some instances, an optional dissolved oxygen probe (pO₂ probe) was placed into the fermentation vessel. The pO₂ probe was calibrated to zero while N₂ was being added to the fermentation vessel and was calibrated to its span (100%) with sterile air, sparging at its initial set point. Tubing used for delivering nutrients, seed culture, extracting solvent, sampling, and base were attached to the head plate and the ends were covered. The fermentation vessel was autoclaved at 121°C for a 30-min liquid cycle.

- 95 -

Propagation Vessel

[00281] The following nutrients were added to the propagation vessel prior to inoculation on a post-inoculation volume basis:

1 kg	15-33% dry corn solids thin mash
1 kg	tap water
30 mg/L	nicotinic acid
30 mg/L	thiamine
0.5 g/L	ethanol
2 g/L	Difco™ yeast extract
1-2 ppm	Lactrol™

[00282] The propagation vessel was inoculated from the seed flask described herein. The shake flask was removed from the incubator/shaker and its contents were centrifuged for 10-15 min at 5000-8000 rpm with a fixed angle rotor between 5-20°C. The supernatant was removed and the wet pellet was re-suspended in < 20% dry corn solids, filter sterilized, thin mash and then was added to the propagation vessel.

Production Vessel

[00283] The following nutrients were added to the production vessel prior to inoculation on a post-inoculation volume basis:

0.5-1.0 kg	25-33% dry corn solids thin mash with or without backset
30 mg/L	nicotinic acid
30 mg/L	thiamine
0.5 g/L	ethanol
2 g/L	urea
1-2 ppm	Lactrol™

[00284] The fermentation broth from the propagation vessel was collected in sterile centrifuge bottles. The material was centrifuged at 5000-8000 rpm for 10 min in a fixed angle rotor between 5-20°C. The supernatant was removed and the wet pellet was re-suspended in < 20% dry corn solids, filter sterilized, thin mash and then was added to the production vessel. Each production vessel received 40-60% of the re-suspended cell pellet.

This process concentrates the cells added to the production vessel. Corn oil fatty acids (0.0-0.7 L/L, post-inoculation volume) were added to the production vessel after inoculation.

[00285] The fermentation vessel (i.e., propagation vessel or production vessel) was operated at 30°C for both propagation and production stages. The pH was allowed to decrease from a pH between 5.4-5.9 to a control set-point of 5.25-5.50 without adding any acid. The pH was controlled for the remainder of the propagation and production stages at a pH = 5.2-5.5 with ammonium hydroxide (propagation) or potassium hydroxide (production). Sterile air was added to the propagation vessel, through the sparger, at 0.2-0.3 slpm for the entire fermentation. Sterile air was added to the production vessel, through the sparger, at 0.2-0.3 slpm for 0-10 hours and then the gas was switched to nitrogen and added to the head space for the remainder of the fermentation. An agitator was used to mix the corn oil fatty acid (i.e., solvent) and aqueous phases. The stir shaft had one to two Rushton impellers below the aqueous level and a third Rushton impeller or marine above the aqueous level. The carbohydrate (glucose) was supplied through simultaneous saccharification and fermentation (SSF) of liquefied corn mash by adding a glucoamylase. The amount of glucose was kept in excess (1-80 g/L) for as long as starch was available for saccharification.

Gas Analysis

[00286] Process air was analyzed on a Thermo Prima db™ (Thermo Fisher Scientific Inc., Waltham, MA) mass spectrometer which was calibrated for these gases: oxygen, nitrogen (balance), helium, carbon dioxide, isobutanol, and argon. The process air was the same process air that was sterilized and then added to each fermentation vessel. The amount of isobutanol stripped, oxygen consumed, and carbon dioxide respired into the off-gas was measured by using the mass spectrometer's mole fraction analysis and gas flow rates (mass flow controller) to the fermentation vessel. The gassing rate per hour was calculated and then that rate was integrated over the course of the fermentation.

Biomass Measurement

[00287] A 5-20 mL sample was removed from a fermentation vessel, placed in a centrifuge tube, and centrifuged. Following centrifugation, the solvent layer (i.e., corn oil fatty acid layer) was removed without removing the layer between the solvent layer and the aqueous

- 97 -

layer. After removal of the solvent layer, the remaining sample was re-suspended by vigorous mixing.

[00288] Cells were diluted by serial dilution for hemacytometer counts. A cover slip was placed on top of the hemacytometer (Hausser Scientific Bright-Line 1492, Horsham, PA). An aliquot (10 μ L) from the final cell dilution was collected by pipette (m20 Variable Channel BioHit pipette with 2-20 μ L BioHit pipette tip, Sartorius Mechatronics Corporation, Bohemia, New York) and injected into the hemacytometer. The hemacytometer was placed on a microscope at 100X-400X magnification for cell counting.

LC Analysis of Fermentation Products in the Aqueous Phase

[00289] Fermentation samples were heated in a heating block at 99°C for 20 min to inactivate the isobutanologen or ethanologen and glucoamylase, and then refrigerated until ready for processing. Samples were removed from refrigeration and allowed to reach room temperature (about one hour). Approximately 300 μ L of a mixed sample was transferred by pipette (m1000 Variable Channel BioHit pipette with 100–1000 μ L BioHit pipette tip, Sartorius Mechatronics Corporation, Bohemia, New York) to a 0.2 μ m centrifuge filter (Nanosep® MF modified nylon centrifuge filter, Pall Corporation, Ann Arbor, MI), then centrifuged for 5 min at 14,000 rpm (Eppendorf 5415C, Eppendorf AG, Hamburg, Germany). Approximately 200 μ L of filtered sample was transferred to a 1.8 autosampler vial with a 250 μ L glass vial insert with polymer feet. A screw cap with PTFE septa was used to cap the vial before vortexing (Vortex-Genie®) the sample at 2700 rpm.

[00290] Samples were analyzed by liquid chromatography (LC) using an Agilent 1200 series LC system equipped with binary, isocratic pumps, vacuum degasser, heated column compartment, sampler cooling system, UV DAD detector, and RI detector (Agilent Technologies, Santa Clara, CA). The column was an Aminex® HPX-87H, 300 X 7.8 with a Bio-Rad Cation H refill, 30X4.6 guard column (Bio-Rad Laboratories, Inc., Hercules, CA). Column temperature was 40°C, with a mobile phase of 0.01 N sulfuric acid at a flow rate of 0.6 mL/min for 40 min.

GC Analysis of Fermentation Products in the Corn Oil Fatty Acid (Solvent) Phase

[00291] Samples were refrigerated until ready for processing. Samples were removed from refrigeration and allowed to reach room temperature (about one hour). Approximately 1000-

2000 μL of sample was transferred using a disposable, bulb pipette to a 1.8 mL autosampler vial. A screw cap with PTFE septa was used to cap the vial.

[00292] Samples were analyzed by gas chromatography (GC) using an Agilent 7890A GC with a 7683B injector and a G2614A auto sampler (Agilent Technologies, Santa Clara, CA). The column was a HP-InnoWax column (30 m x 0.32 mm ID, 0.25 μm film).

Samples

[00293] Samples are described in Table 9. Results for the isobutanologen are shown in Figures 11A-11D, and the results for the ethanologen are shown in Figures 12A-12D. TCER is total carbon dioxide evolution rate (mmol CO_2 produced per hour); biomass is cfu/mL; production rate is g/L/h, aqueous phase; and glucose equivalents consumed is g/L.

Table 9

Sample	Microorganism	Backset (% water volume)
A	Isobutanologen	0
B	Isobutanologen	15%
C	Isobutanologen	30%
D	Ethanologen	0
E	Ethanologen	30%

[00294] Figure 11A demonstrates CO_2 evolution rates (mmol(s) per hour) with an isobutanologen with backset and without backset. Figure 11B demonstrates isobutanologen biomass concentrations as cell counts with backset and without backset. Figure 11C demonstrates isobutanol volumetric productivity (grams per liter per hour) with backset and without backset. Figure 11D demonstrates glucose equivalent consumption rates (grams per liter per hour) with an isobutanologen with backset and without backset.

[00295] Figure 12A demonstrates CO_2 evolution rates (mmol(s) per hour) with an ethanologen with backset and without backset. Figure 12B demonstrates ethanologen biomass concentrations as cell counts with backset and without backset. Figure 12C demonstrates ethanol volumetric productivity (grams per liter per hour) with backset and without backset. Figure 12D demonstrates glucose equivalent consumption rates (grams per liter per hour) with an ethanologen with backset and without backset.

- 99 -

[00296] These experiments show that when backset is added to the liquefaction step of an isobutanologen fermentation, the volumetric productivity of isobutanol is improved as compared to an isobutanologen fermentation in the absence of backset. In addition, the improvement in the volumetric productivity of an isobutanologen fermentation was greater than the benefit shown in an ethanologen process.

[00297] All documents cited herein, including journal articles or abstracts, published or corresponding U.S. or foreign patent applications, issued or foreign patents, or any other documents, are each entirely incorporated by reference herein, including all data, tables, figures, and text presented in the cited documents.

[00298] Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.

EDITORIAL NOTE

Case No. 2013323396

Please notionally renumber the Claim numbers between Claim numbers 4 and 9, to Claim numbers 5 – 8.

The Claims defining the invention are as follows:

1. A method for producing butanol comprising:
 - a) providing a recombinant host cell comprising a butanol biosynthetic pathway; and
 - b) contacting the recombinant host cell with a fermentation medium comprising:
 - i) a fermentable carbon substrate,
 - ii) magnesium, and
 - iii) backset wherein the backset is added to the fermentation medium as at least 5% of the water volume of the fermentation medium;wherein butanol is produced via the butanol biosynthetic pathway.
2. The method of claim 1, wherein magnesium is added during propagation of the recombinant host cell.
3. The method of claim 1 or 2, wherein magnesium or a portion thereof is added as a magnesium salt or a concentrated magnesium salt solution.
4. The method of claim 1, wherein the magnesium in the fermentation medium is in the range of about 5 mM to about 200 mM.
6. The method of claim 1, wherein the magnesium in the fermentation medium is in the range of about 10 mM to about 150 mM.
7. The method of claim 1, wherein the magnesium in the fermentation medium is in the range of about 30 mM to about 70 mM.
8. The method of claim 1, wherein the magnesium in the fermentation medium is in the range of about 50 mM to about 150 mM.
8. The method of any one of claims 1 to 7, wherein the butanol is isobutanol.
9. The method of any one of claims 1 to 7, wherein the butanol biosynthetic pathway is an isobutanol biosynthetic pathway.
10. The method of claim 9, wherein the isobutanol biosynthetic pathway comprises the following substrate to product conversions:

- i) pyruvate to acetolactate;
 - ii) acetolactate to 2,3-dihydroxyisovalerate;
 - iii) 2,3-dihydroxyisovalerate to α -ketoisovalerate;
 - iv) α -ketoisovalerate to isobutyraldehyde; and
 - v) isobutyraldehyde to isobutanol.
11. The method of claim 10, wherein the isobutanol biosynthetic pathway comprises polynucleotides encoding polypeptides having acetolactate synthase, keto acid reductoisomerase, dihydroxy acid dehydratase, ketoisovalerate decarboxylase, and alcohol dehydrogenase activity.
12. The method of any one of claims 1 to 11, wherein the recombinant host cell is selected from bacteria, cyanobacteria, filamentous fungi, and yeast.
13. The method of claim 12, wherein the recombinant host cell is selected from *Clostridium*, *Zymomonas*, *Escherichia*, *Salmonella*, *Serratia*, *Erwinia*, *Klebsiella*, *Shigella*, *Rhodococcus*, *Pseudomonas*, *Bacillus*, *Lactobacillus*, *Enterococcus*, *Alcaligenes*, *Klebsiella*, *Paenibacillus*, *Arthrobacter*, *Corynebacterium*, *Brevibacterium*, *Schizosaccharomyces*, *Kluyveromyces*, *Yarrowia*, *Pichia*, *Zygosaccharomyces*, *Debaryomyces*, *Candida*, *Brettanomyces*, *Pachysolen*, *Hansenula*, *Issatchenkia*, *Trichosporon*, *Yamadazyma*, and *Saccharomyces*.

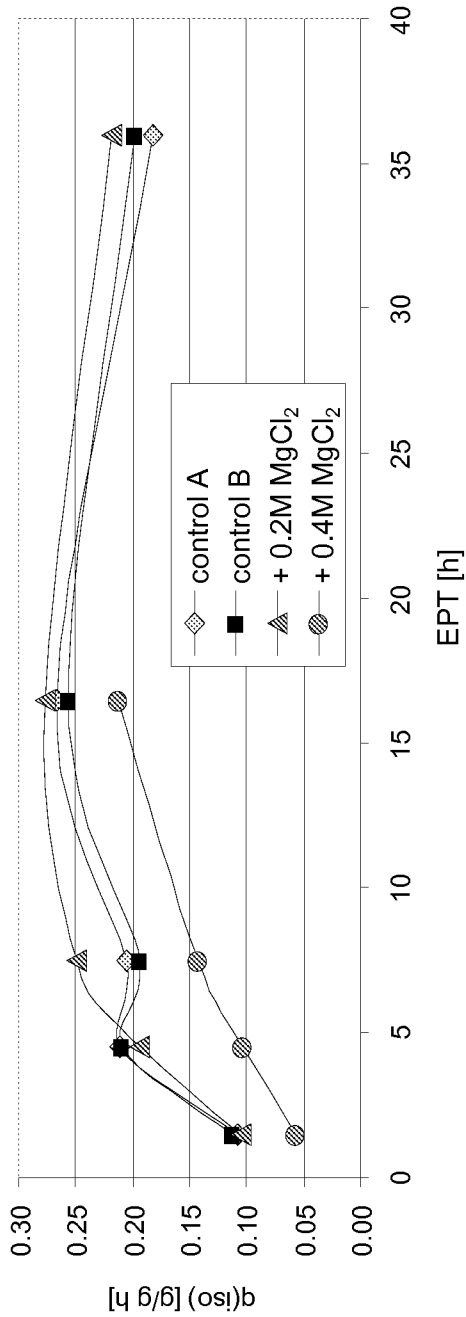


FIG. 1

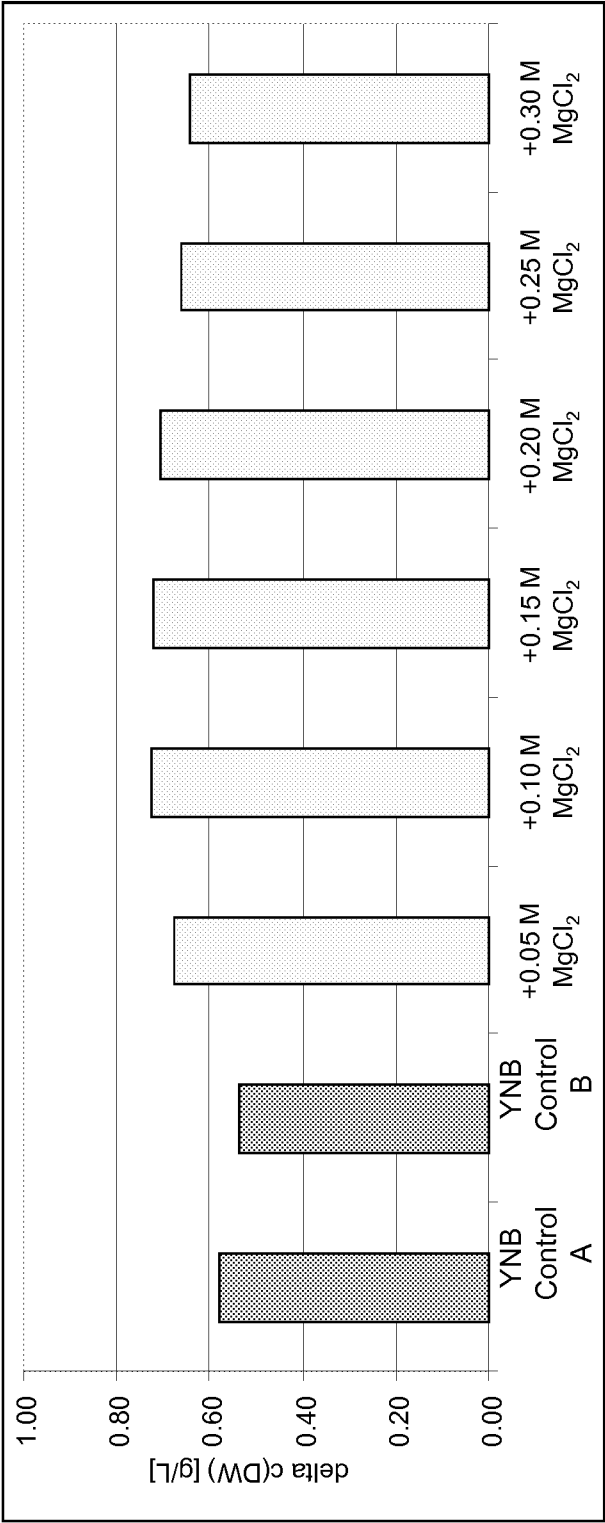


FIG. 2

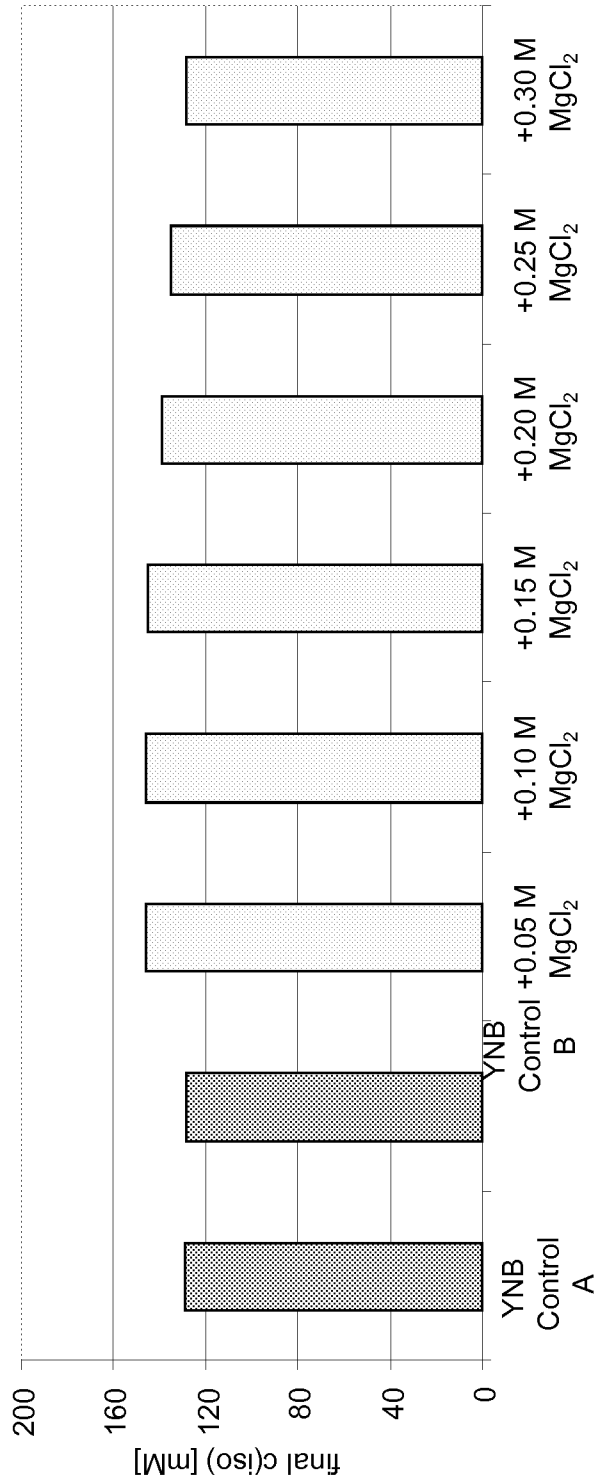


FIG. 3

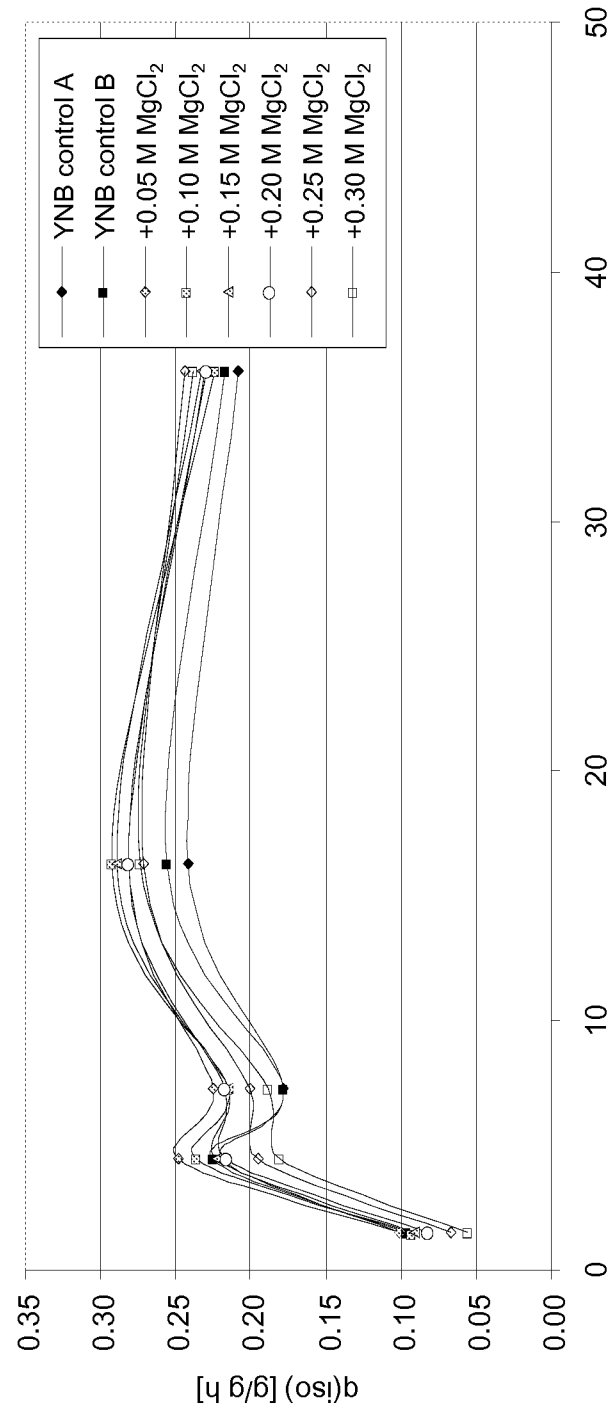


FIG. 4

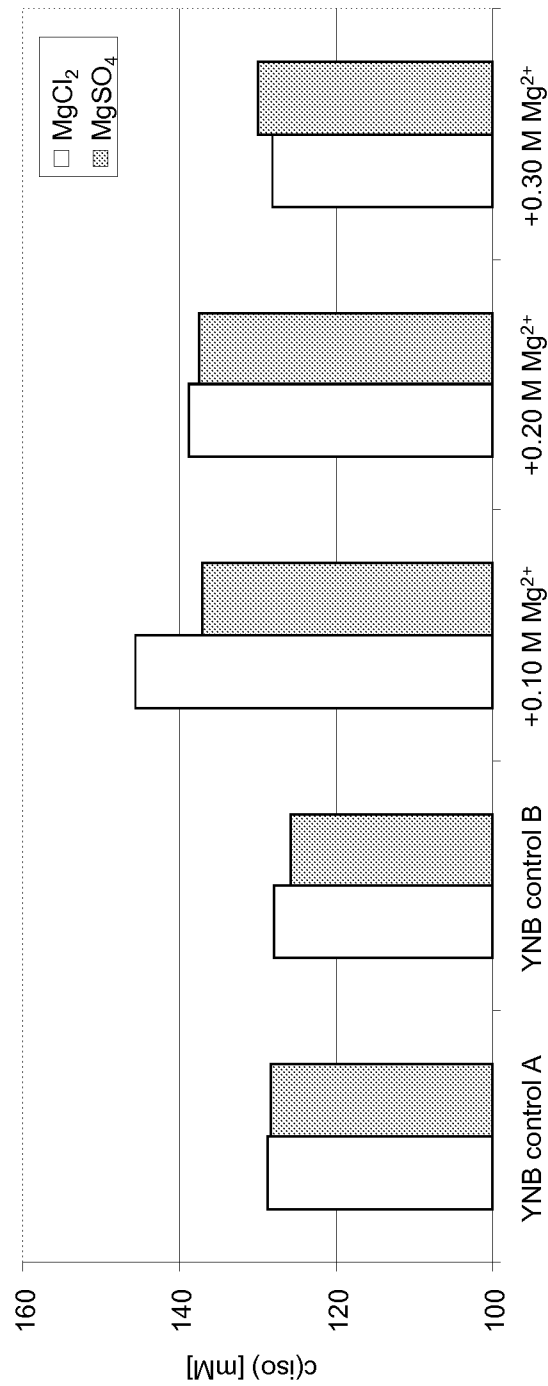


FIG. 5

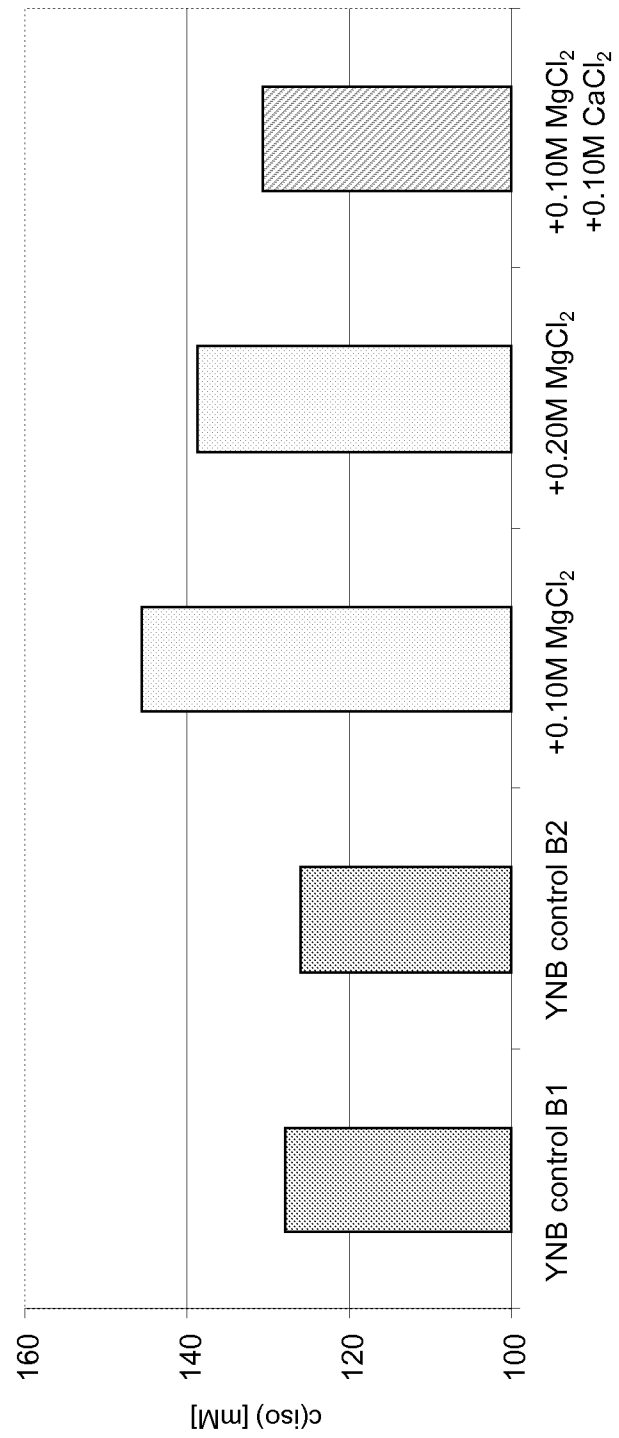


FIG. 6

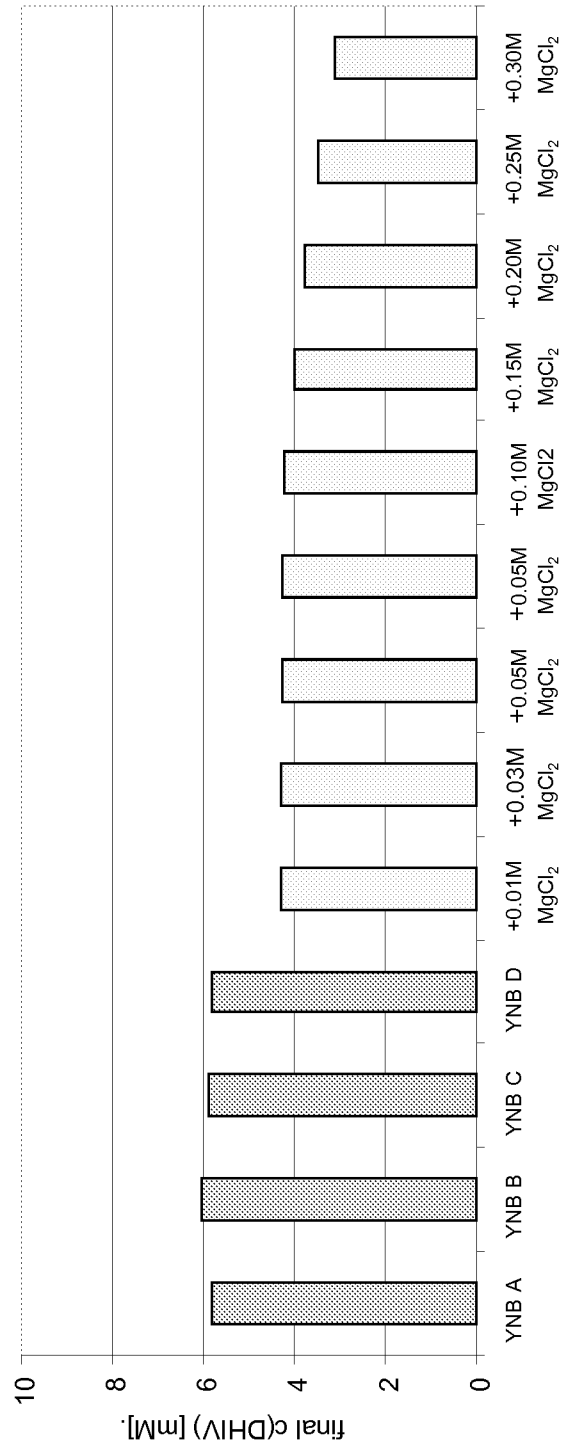


FIG. 7

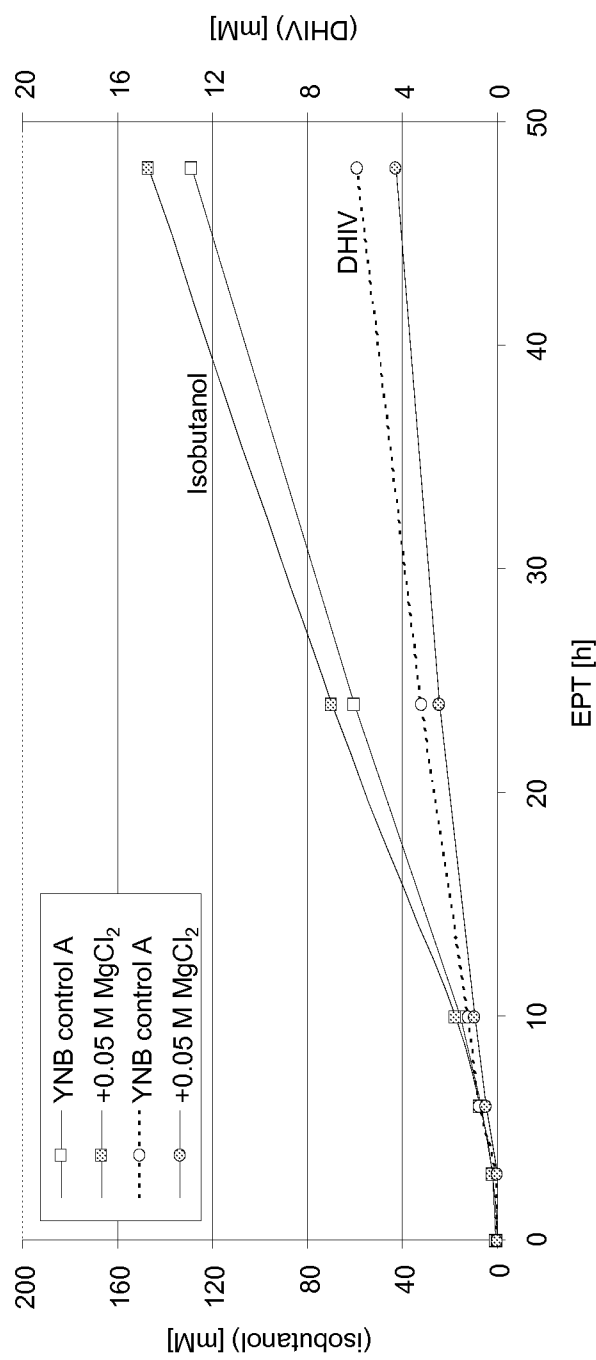


FIG. 8

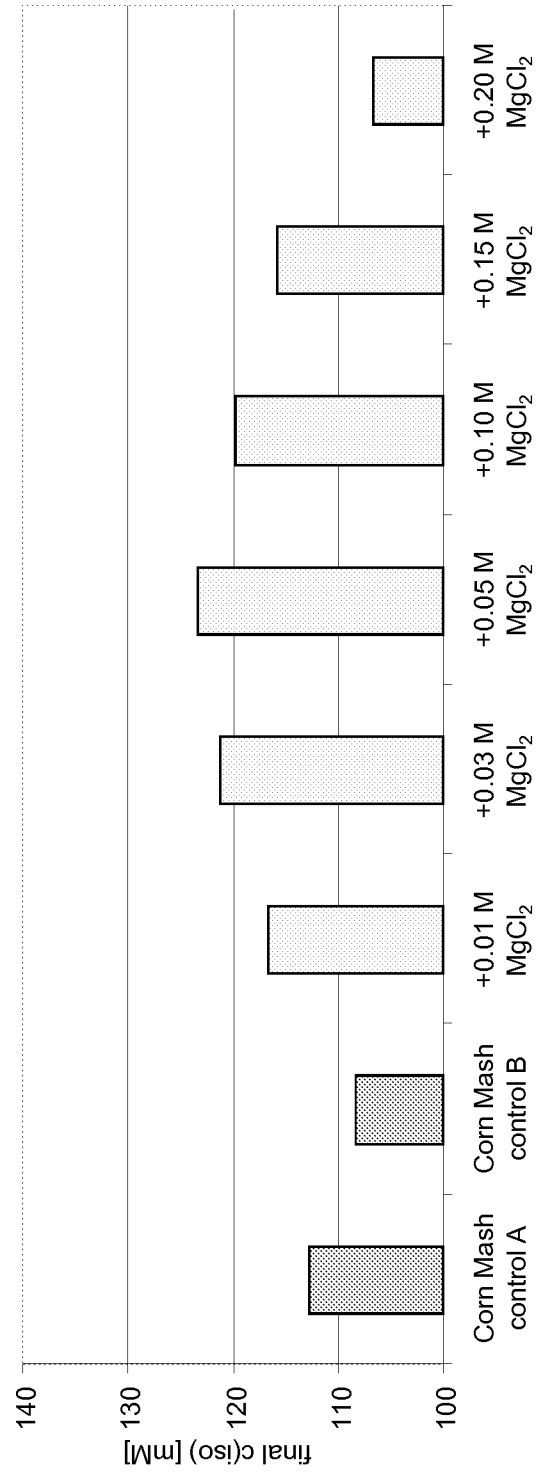


FIG. 9

10/18

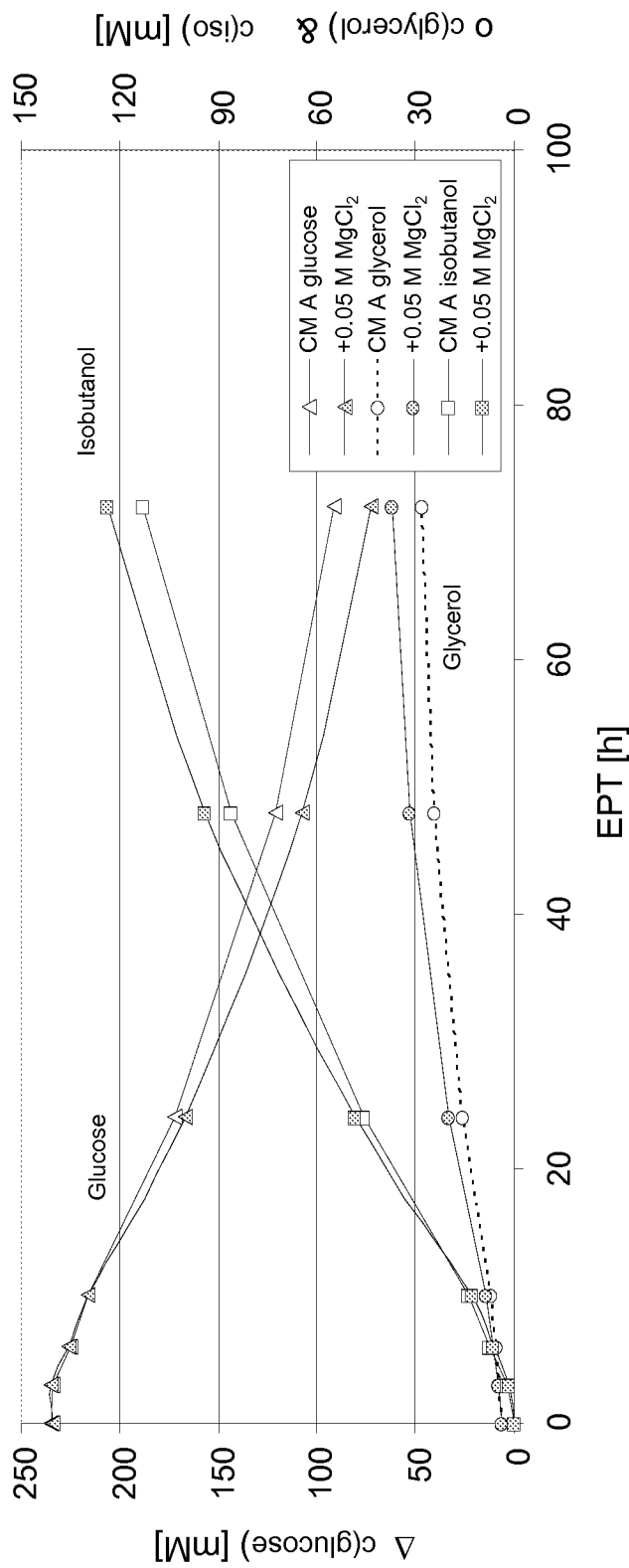


FIG. 10

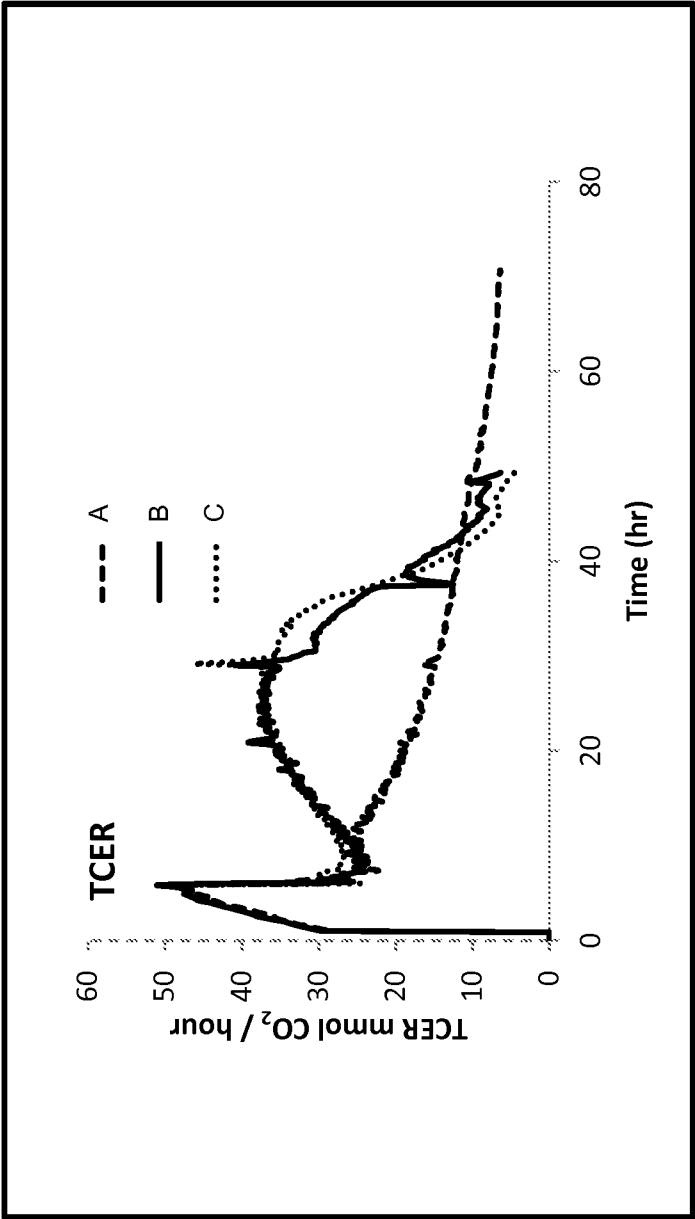


FIG. 11A

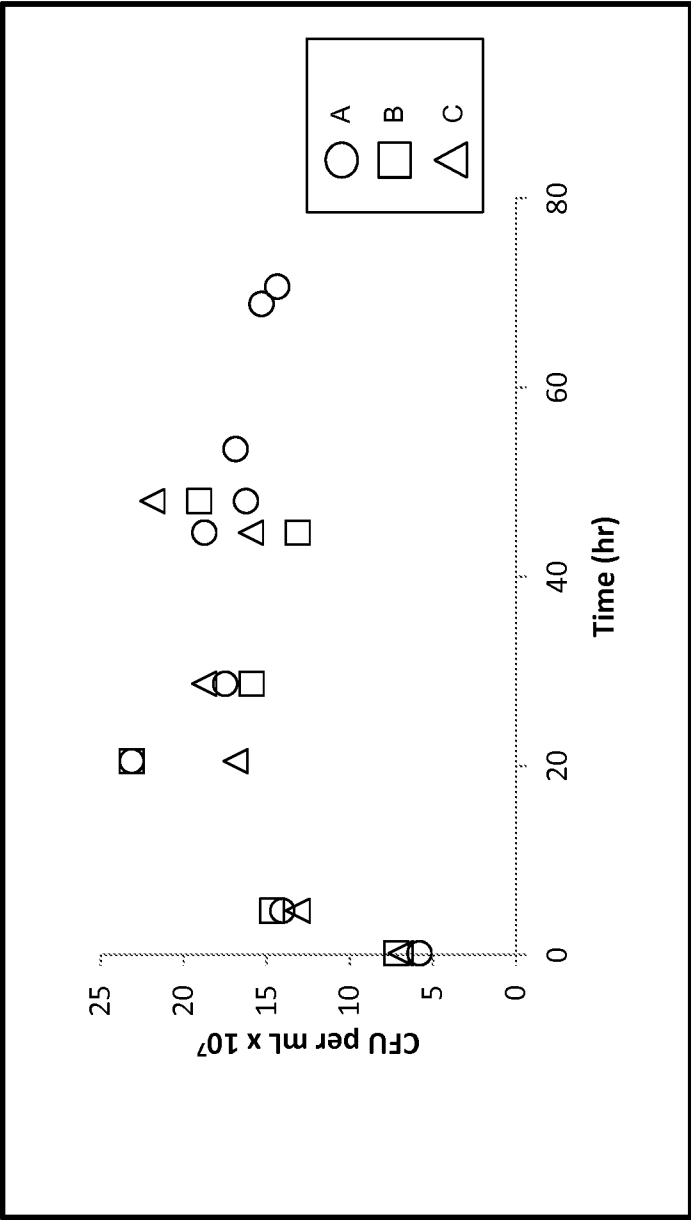


FIG. 11B

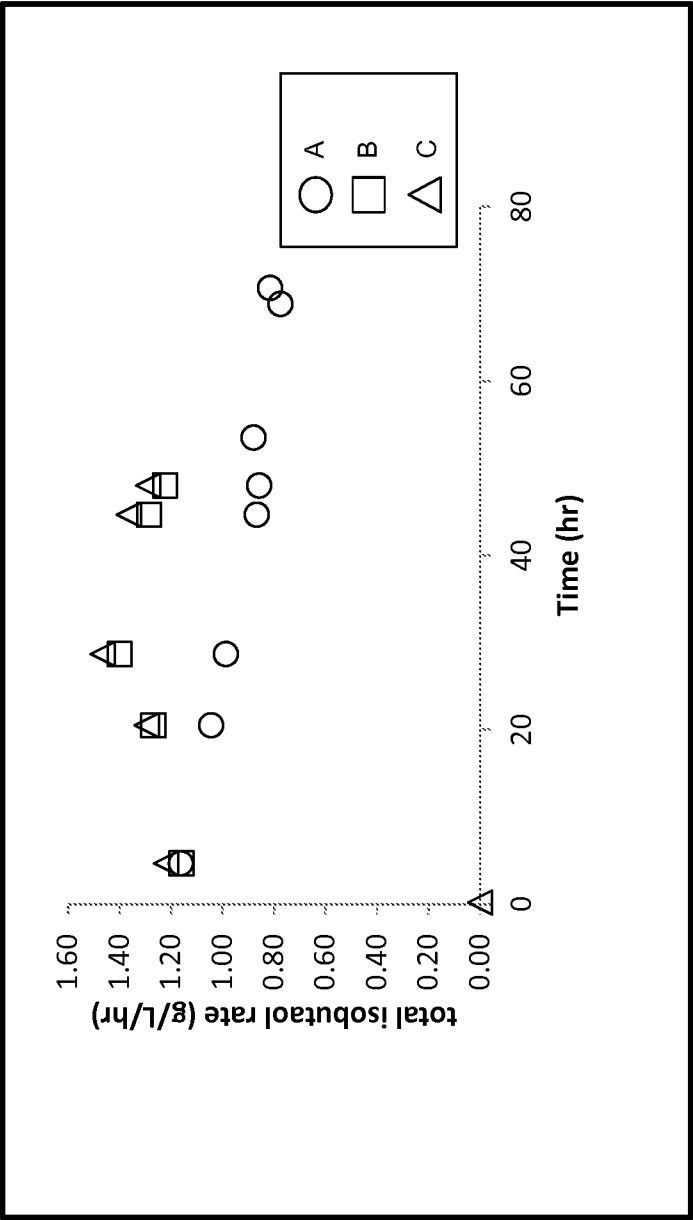


FIG. 11C

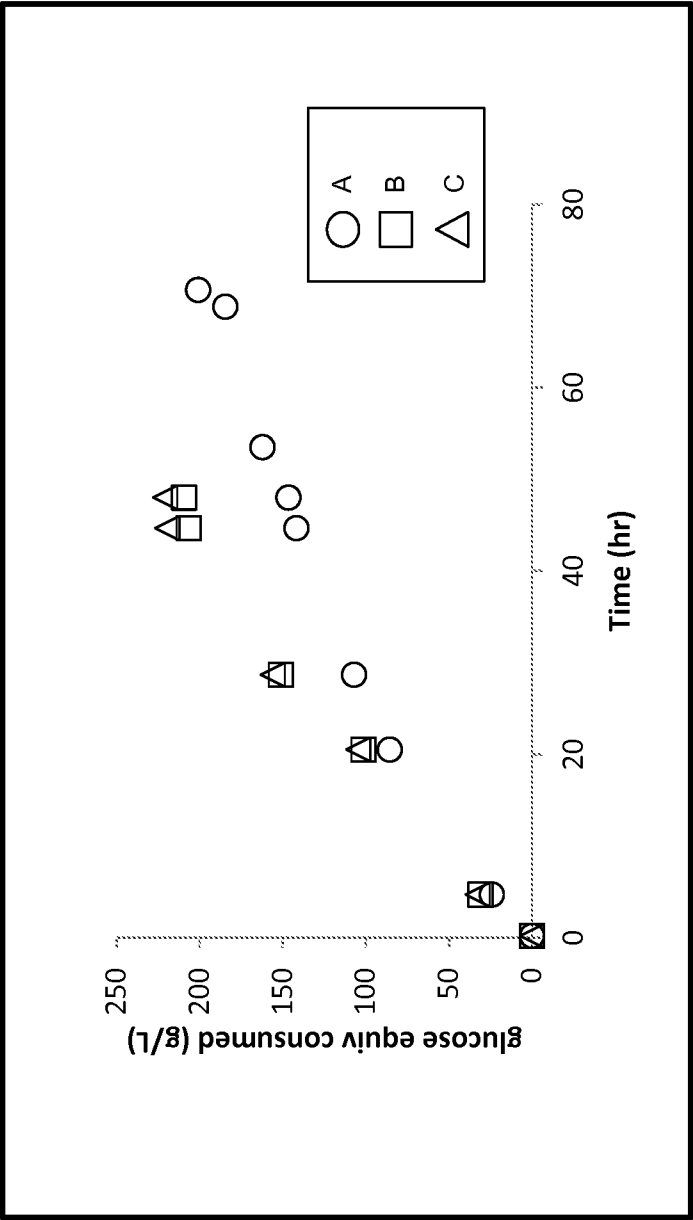


FIG. 11D

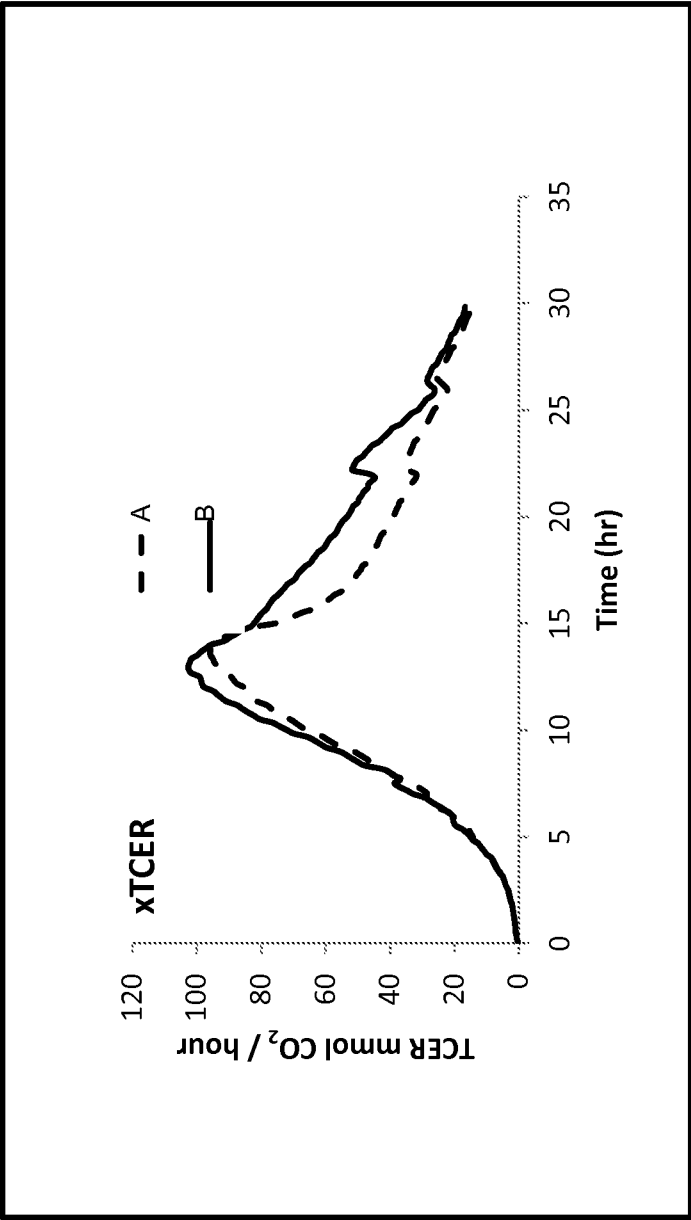


FIG. 12A

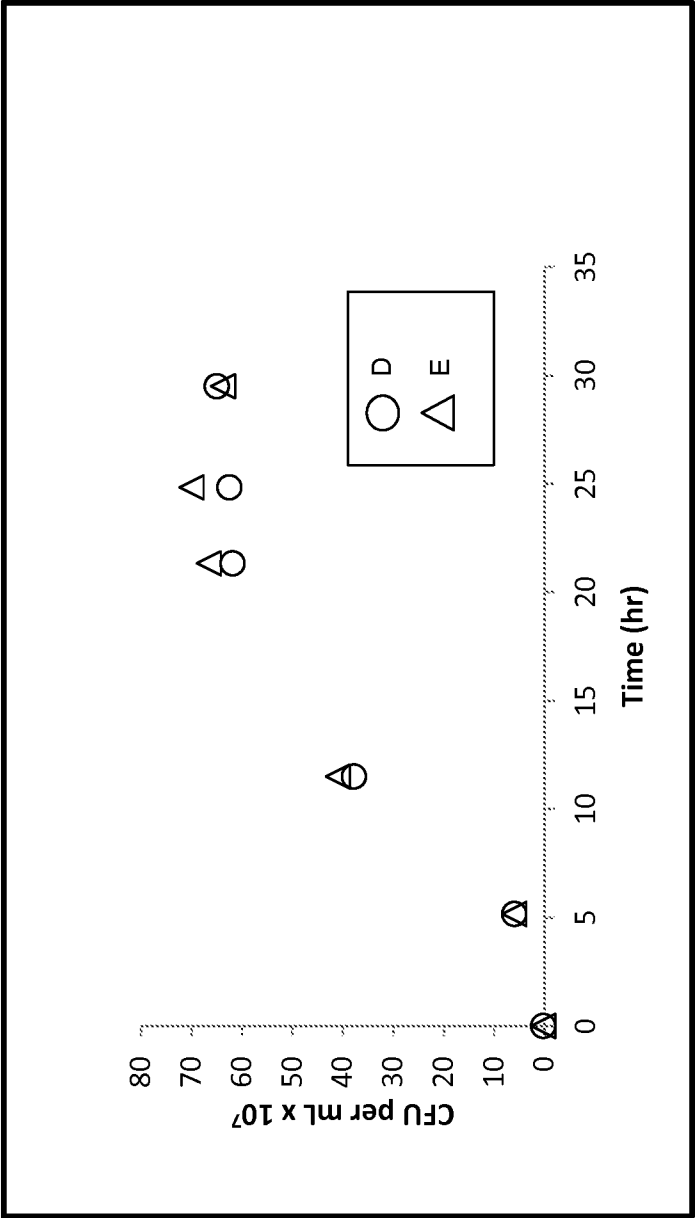


FIG. 12B

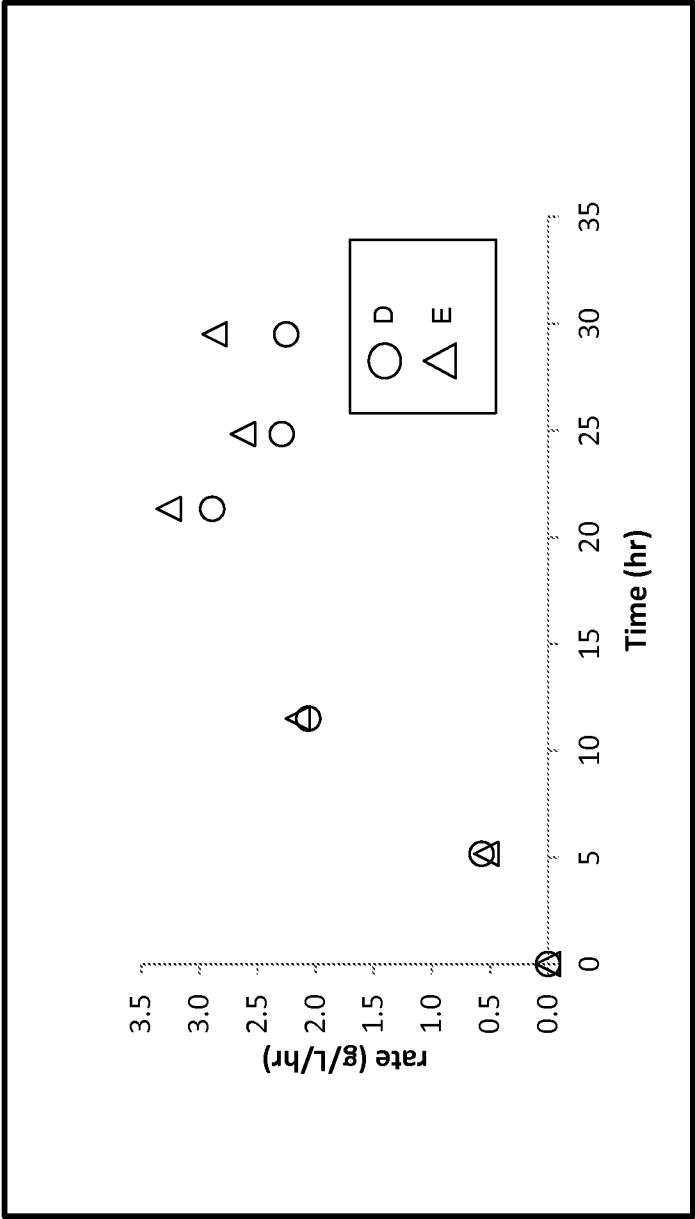


FIG. 12C

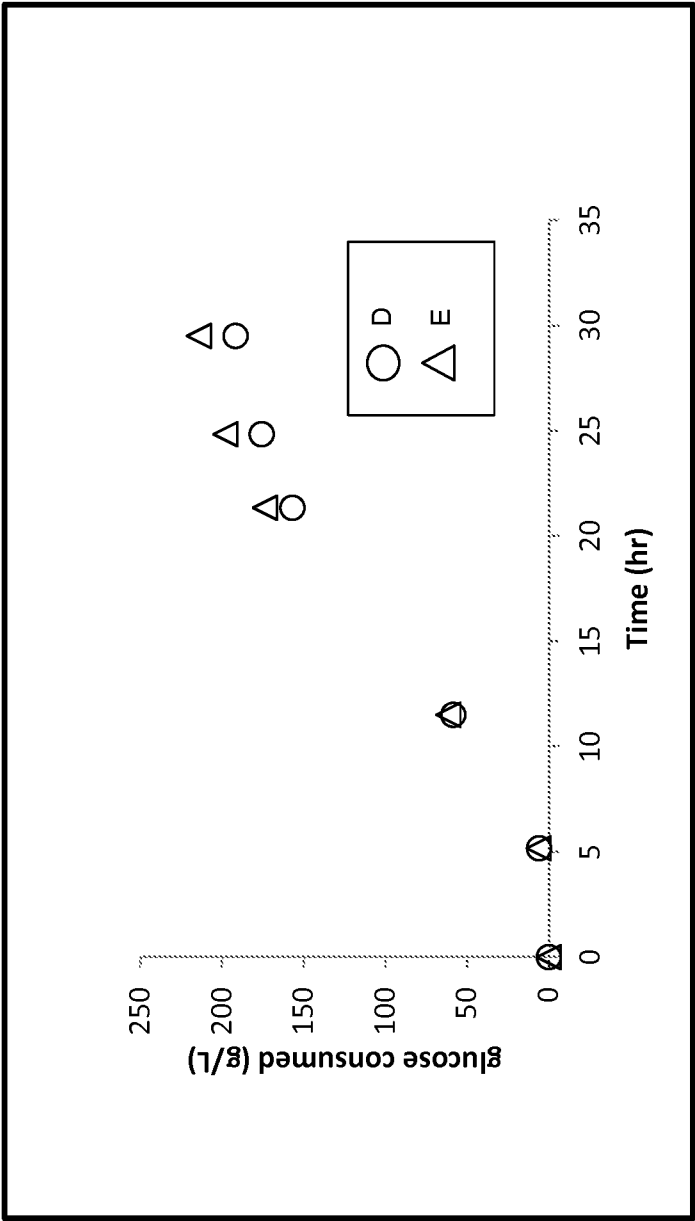


FIG. 12D

20130927_CL5646WOPCT_ST25.txt
SEQUENCE LISTING

<110> Butamax Advanced Biofuels LLC

<120> PRODUCTION OF FERMENTATION PRODUCTS

<130> CL5646WOPCT

<160> 201

<170> PatentIn version 3.5

<210> 1

<211> 570

<212> PRT

<213> Bacillus subtilis

<400> 1

Met Thr Lys Ala Thr Lys Glu Gln Lys Ser Leu Val Lys Asn Arg Gly
1 5 10 15

Ala Glu Leu Val Val Asp Cys Leu Val Glu Gln Gly Val Thr His Val
20 25 30

Phe Gly Ile Pro Gly Ala Lys Ile Asp Ala Val Phe Asp Ala Leu Gln
35 40 45

Asp Lys Gly Pro Glu Ile Ile Val Ala Arg His Glu Gln Asn Ala Ala
50 55 60

Phe Met Ala Gln Ala Val Gly Arg Leu Thr Gly Lys Pro Gly Val Val
65 70 75 80

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

Thr Ala Asn Thr Glu Gly Asp Pro Val Val Ala Leu Ala Gly Asn Val
100 105 110

Ile Arg Ala Asp Arg Leu Lys Arg Thr His Gln Ser Leu Asp Asn Ala
115 120 125

Ala Leu Phe Gln Pro Ile Thr Lys Tyr Ser Val Glu Val Gln Asp Val
130 135 140

Lys Asn Ile Pro Glu Ala Val Thr Asn Ala Phe Arg Ile Ala Ser Ala
145 150 155 160

Gly Gln Ala Gly Ala Ala Phe Val Ser Phe Pro Gln Asp Val Val Asn
165 170 175

Glu Val Thr Asn Thr Lys Asn Val Arg Ala Val Ala Ala Pro Lys Leu
180 185 190

Gly Pro Ala Ala Asp Asp Ala Ile Ser Ala Ala Ile Ala Lys Ile Gln

195

200

205

Thr Ala Lys Leu Pro Val Val Leu Val Gly Met Lys Gly Gly Arg Pro
 210 215 220

Glu Ala Ile Lys Ala Val Arg Lys Leu Leu Lys Lys Val Gln Leu Pro
 225 230 235 240

Phe Val Glu Thr Tyr Gln Ala Ala Gly Thr Leu Ser Arg Asp Leu Glu
 245 250 255

Asp Gln Tyr Phe Gly Arg Ile Gly Leu Phe Arg Asn Gln Pro Gly Asp
 260 265 270

Leu Leu Leu Glu Gln Ala Asp Val Val Leu Thr Ile Gly Tyr Asp Pro
 275 280 285

Ile Glu Tyr Asp Pro Lys Phe Trp Asn Ile Asn Gly Asp Arg Thr Ile
 290 295 300

Ile His Leu Asp Glu Ile Ile Ala Asp Ile Asp His Ala Tyr Gln Pro
 305 310 315 320

Asp Leu Glu Leu Ile Gly Asp Ile Pro Ser Thr Ile Asn His Ile Glu
 325 330 335

His Asp Ala Val Lys Val Glu Phe Ala Glu Arg Glu Gln Lys Ile Leu
 340 345 350

Ser Asp Leu Lys Gln Tyr Met His Glu Gly Glu Gln Val Pro Ala Asp
 355 360 365

Trp Lys Ser Asp Arg Ala His Pro Leu Glu Ile Val Lys Glu Leu Arg
 370 375 380

Asn Ala Val Asp Asp His Val Thr Val Thr Cys Asp Ile Gly Ser His
 385 390 395 400

Ala Ile Trp Met Ser Arg Tyr Phe Arg Ser Tyr Glu Pro Leu Thr Leu
 405 410 415

Met Ile Ser Asn Gly Met Gln Thr Leu Gly Val Ala Leu Pro Trp Ala
 420 425 430

Ile Gly Ala Ser Leu Val Lys Pro Gly Glu Lys Val Val Ser Val Ser
 435 440 445

Gly Asp Gly Gly Phe Leu Phe Ser Ala Met Glu Leu Glu Thr Ala Val
 450 455 460

Arg Leu Lys Ala Pro Ile Val His Ile Val Trp Asn Asp Ser Thr Tyr

465 470 475 480

Asp Met Val Ala Phe Gln Gln Leu Lys Lys Tyr Asn Arg Thr Ser Ala
485 490 495

Val Asp Phe Gly Asn Ile Asp Ile Val Lys Tyr Ala Glu Ser Phe Gly
500 505 510

Ala Thr Gly Leu Arg Val Glu Ser Pro Asp Gln Leu Ala Asp Val Leu
515 520 525

Arg Gln Gly Met Asn Ala Glu Gly Pro Val Ile Ile Asp Val Pro Val
530 535 540

Asp Tyr Ser Asp Asn Ile Asn Leu Ala Ser Asp Lys Leu Pro Lys Glu
545 550 555 560

Phe Gly Glu Leu Met Lys Thr Lys Ala Leu
565 570

<210> 2
<211> 1716
<212> DNA
<213> Bacillus subtilis

<400> 2
atgttgacaa aagcaacaaa agaacaaaaa tcccttgtga aaaacagagg ggcggagctt 60
gttgttgatt gcttagtgga gcaaggtgtc acacatgtat ttggcattcc aggtgcaaaa 120
attgatgcgg tatttgacgc ttacaagat aaaggacctg aaattatcgt tgcccggcac 180
gaacaaaacg cagcattcat ggccaagca gtcggccgtt taactggaaa accgggagtc 240
gtgttagtca catcaggacc ggggtgcctct aacttggcaa caggcctgct gacagcgaac 300
actgaaggag accctgtcgt tgcgcttgct ggaaacgtga tccgtgcaga tcgtttaaaa 360
cggacacatc aatctttgga taatgcggcg ctattccagc cgattacaaa atacagtgtg 420
gaagttcaag atgtaaaaaa tataccggaa gctgttaca atgcatttag gatagcgtca 480
gcagggcagg ctggggccgc ttttgtgagc tttccgaag atgttgtgaa tgaagtcaca 540
aatacgaaaa acgtgcgtgc tgttgacgcg caaaactcg gtcctgcagc agatgatgca 600
atcagtgcgg ccatagcaaa aatccaaaca gcaaaacttc ctgtcgtttt ggtcggcatg 660
aaaggcggaa gaccggaagc aattaaagcg gttcgcaagc ttttgaaaaa ggttcagctt 720
ccatttgttg aaacatatca agctgccggg accctttcta gagatttaga ggatcaatat 780
tttggccgta tcggtttgtt ccgcaaccag cctggcgatt tactgctaga gcaggcagat 840
gttgttctga cgatcggcta tgacccgatt gaatatgatc cgaaattctg gaatatcaat 900
ggagaccgga caattatcca tttagacgag attatcgctg acattgatca tgcttaccag 960
cctgatcttg aattgatcgg tgacattccg tccacgatca atcatatcga acacgatgct 1020
gtgaaagtgg aatttgacga gcgtgagcag aaaatccttt ctgattttaa acaatatatg 1080

```

catgaagggtg agcaggtgcc tgcagattgg aaatcagaca gagcgcaccc tcttgaaatc      1140
gttaaagagt tgcgtaatgc agtcgatgat catgttacag taacttgcga tatcggttcg      1200
cacgccattt ggatgtcacg ttatttccgc agctacgagc cgttaacatt aatgatcagt      1260
aacggtatgc aaacactcgg cgttgcgctt ccttgggcaa tcggcgcttc attggtgaaa      1320
ccggggagaaa aagtggtttc tgtctctggt gacggcggtt tcttattctc agcaatggaa      1380
ttagagacag cagttcgact aaaagcacca attgtacaca ttgtatggaa cgacagcaca      1440
tatgacatgg ttgcattcca gcaattgaaa aaatataacc gtacatctgc ggtcgaatttc      1500
ggaaatatcg atatcgtgaa atatgcggaa agcttcggag caactggctt gcgcgtagaa      1560
tcaccagacc agctggcaga tgttctgcgt caaggcatga acgctgaagg tcctgtcatc      1620
atcgatgtcc cggttgacta cagtataac attaatttag caagtgacaa gcttccgaaa      1680
gaattcgggg aactcatgaa aacgaaagct ctctag                                  1716

```

```

<210> 3
<211> 559
<212> PRT
<213> Klebsiella pneumoniae

```

```
<400> 3
```

```
Met Asp Lys Gln Tyr Pro Val Arg Gln Trp Ala His Gly Ala Asp Leu
1          5          10          15

```

```
Val Val Ser Gln Leu Glu Ala Gln Gly Val Arg Gln Val Phe Gly Ile
          20          25          30

```

```
Pro Gly Ala Lys Ile Asp Lys Val Phe Asp Ser Leu Leu Asp Ser Ser
          35          40          45

```

```
Ile Arg Ile Ile Pro Val Arg His Glu Ala Asn Ala Ala Phe Met Ala
          50          55          60

```

```
Ala Ala Val Gly Arg Ile Thr Gly Lys Ala Gly Val Ala Leu Val Thr
          65          70          75          80

```

```
Ser Gly Pro Gly Cys Ser Asn Leu Ile Thr Gly Met Ala Thr Ala Asn
          85          90          95

```

```
Ser Glu Gly Asp Pro Val Val Ala Leu Gly Gly Ala Val Lys Arg Ala
          100          105          110

```

```
Asp Lys Ala Lys Gln Val His Gln Ser Met Asp Thr Val Ala Met Phe
          115          120          125

```

```
Ser Pro Val Thr Lys Tyr Ala Ile Glu Val Thr Ala Pro Asp Ala Leu
          130          135          140

```

```
Ala Glu Val Val Ser Asn Ala Phe Arg Ala Ala Glu Gln Gly Arg Pro

```

145		150		155		160
Gly	Ser	Ala	Phe	Val ₁₆₅	Ser	Leu
					Pro	Gln
				Asp ₁₇₀	Val	Val
					Asp	Gly
					Pro	Val ₁₇₅
Ser	Gly	Lys	Val ₁₈₀	Leu	Pro	Ala
					Ser	Gly
					Ala	Pro
					Gln	Met
					Gly ₁₉₀	Ala
					Ala	Ala
Pro	Asp	Asp ₁₉₅	Ala	Ile	Asp	Gln
					Val ₂₀₀	Ala
					Lys	Leu
					Ile	Ala
					Gln ₂₀₅	Ala
					Lys	
Asn	Pro ₂₁₀	Ile	Phe	Leu	Leu	Gly
					Leu	Met
					Ala	Ser
					Gln ₂₂₀	Pro
					Glu	Asn
					Ser	
Lys	Ala	Leu	Arg	Arg	Leu	Leu
₂₂₅					Glu	Thr
					Ser	His
					Ile ₂₃₅	Pro
					Val	Thr
					Ser	₂₄₀
Thr	Tyr	Gln	Ala	Ala ₂₄₅	Gly	Ala
					Val	Asn
					Gln ₂₅₀	Asp
					Asn	Phe
					Ser	Arg
					₂₅₅	Phe
Ala	Gly	Arg	Val ₂₆₀	Gly	Leu	Phe
					Asn	Asn
					Gln ₂₆₅	Ala
					Gly	Asp
					Arg ₂₇₀	Leu
					Leu	
Gln	Leu	Ala ₂₇₅	Asp	Leu	Val	Ile
					Cys ₂₈₀	Ile
					Gly	Tyr
					Ser	Pro
					Val ₂₈₅	Glu
					Tyr	
Glu	Pro ₂₉₀	Ala	Met	Trp	Asn	Ser
					Gly ₂₉₅	Asn
					Ala	Thr
					Leu ₃₀₀	Val
					His	Ile
					Asp	
Val	Leu	Pro	Ala	Tyr	Glu	Glu
₃₀₅					Arg ₃₁₀	Asn
					Tyr	Thr
					Pro ₃₁₅	Asp
					Val	Glu
					Leu ₃₂₀	
Val	Gly	Asp	Ile	Ala ₃₂₅	Gly	Thr
					Leu	Asn
					Lys ₃₃₀	Leu
					Ala	Gln
					Asn ₃₃₅	Ile
					Asp	
His	Arg	Leu	Val ₃₄₀	Leu	Ser	Pro
					Gln ₃₄₅	Ala
					Ala	Glu
					Ile	Leu
					Arg ₃₅₀	Asp
					Arg	
Gln	His	Gln ₃₅₅	Arg	Glu	Leu	Leu
					Asp ₃₆₀	Arg
					Arg	Gly
					Ala	Gln
					Leu ₃₆₅	Asn
					Gln	
Phe	Ala ₃₇₀	Leu	His	Pro	Leu	Arg
					Ile ₃₇₅	Val
					Arg	Ala
					Met ₃₈₀	Gln
					Asp	Ile
					Val	
Asn	Ser	Asp	Val	Thr	Leu	Thr
₃₈₅					Val ₃₉₀	Val
					Asp	Met
					Gly ₃₉₅	Ser
					Phe	His
					Ile	Trp
					₄₀₀	
Ile	Ala	Arg	Tyr	Leu	Tyr	Thr
					Phe	Arg
					Ala ₄₁₀	Arg
					Gln	Val
					Met	Ile
					₄₁₅	Ser
Asn	Gly	Gln	Gln	Thr	Met	Gly
					Val	Ala
					Leu	Pro
					Trp	Ala
					Ile	Gly
					Ala	

420

425

430

Trp Leu Val Asn Pro Glu Arg Lys Val Val Ser Val Ser Gly Asp Gly
 435 440 445

Gly Phe Leu Gln Ser Ser Met Glu Leu Glu Thr Ala Val Arg Leu Lys
 450 455 460

Ala Asn Val Leu His Leu Ile Trp Val Asp Asn Gly Tyr Asn Met Val
 465 470 475 480

Ala Ile Gln Glu Glu Lys Lys Tyr Gln Arg Leu Ser Gly Val Glu Phe
 485 490 495

Gly Pro Met Asp Phe Lys Ala Tyr Ala Glu Ser Phe Gly Ala Lys Gly
 500 505 510

Phe Ala Val Glu Ser Ala Glu Ala Leu Glu Pro Thr Leu Arg Ala Ala
 515 520 525

Met Asp Val Asp Gly Pro Ala Val Val Ala Ile Pro Val Asp Tyr Arg
 530 535 540

Asp Asn Pro Leu Leu Met Gly Gln Leu His Leu Ser Gln Ile Leu
 545 550 555

<210> 4
 <211> 2055
 <212> DNA
 <213> Klebsiella pneumoniae

<400> 4
 tcgaccacgg ggtgctgacc ttcggcgaaa ttcacaagct gatgatcgac ctgcccgcgc 60
 acagcgcgtt cctgcaggct aatctgcata ccgataatct cgatgccgcc atccgttccg 120
 tagaaagtta aggggggtcac atggacaaac agtatccggt acgccagtgg gcgcacggcg 180
 ccgatctcgt cgtcagtcag ctggaagctc agggagtacg ccaggtgttc ggcatccccg 240
 gcgcacaaaat cgacaaggctc tttgattcac tgctggattc ctccattcgc attattccgg 300
 tacgccacga agccaacgcc gcatttatgg ccgccgccgt cggacgcatt accggcaaag 360
 cgggcgtggc gctggtcacc tccggtccgg gctgttccaa cctgatcacc ggcatggcca 420
 ccgcgaacag cgaaggcgac ccggtggtgg ccctgggcgg cgcggtaaaa cgcgccgata 480
 aagcgaagca ggtccaccag agtatggata cggtggcgat gttcagcccg gtcaccaa 540
 acgcatcga ggtgacggcg ccgatgcgc tggcggaagt ggtctccaac gccttccgcg 600
 ccgccgagca gggccggccg ggcagcgcgt tcgttagcct gccgcaggat gtggtcgatg 660
 gccgggtcag cggcaaagtg ctgccggcca gcggggcccc gcagatgggc gccgcgccgg 720
 atgatgcat cgaccaggtg gcgaagctta tcgccaggc gaagaaccg atcttctgc 780
 tcggcctgat ggccagccag ccggaaaaca gcaaggcgct gcgccgtttg ctggagacca 840

20130927_CL5646WOPCT_ST25. txt

gccatattcc agtcaccagc acctatcagg ccgccggagc ggtgaatcag gataacttct 900
ctcgtttcgc cggccgggtt gggctgttta acaaccaggc cggggaccgt ctgctgcagc 960
tcgccgacct ggtgatctgc atcgggtaca gcccggtgga atacgaaccg gcgatgtgga 1020
acagcggcaa cgcgacgtg gtgcacatcg acgtgctgcc cgcctatgaa gagcgcaact 1080
acaccccgga tgtcgagctg gtgggcgata tcgccggcac tctcaacaag ctggcgcaaa 1140
atatcgatca tcggctggtg ctctccccgc aggcggcgga gatcctccgc gaccgccagc 1200
accagcgca gctgctggac cggcgggcg cgcagctcaa ccagtttgcc ctgcatcccc 1260
tgcgcatcgt tcgcgccatg caggatatcg tcaacagcga cgtcacgttg accgtggaca 1320
tgggcagctt ccatatctgg attgcccgt acctgtacac gttccgcgcc cgtcaggtga 1380
tgatctcaa cggccagcag accatggcg tcgccctgcc ctgggctatc ggcgcctggc 1440
tggtcaatcc tgagcgcaa gtggtctccg tctccggcga cggcggttc ctgcagtcga 1500
gcatggagct ggagaccgcc gtccgcctga aagccaacgt gctgcatctt atctgggtcg 1560
ataacggcta caacatggc gctatccagg aagagaaaa atatcagcgc ctgtccggcg 1620
tcgagtttg gccgatggat tttaaagcct atgccgaatc cttcggcgcg aaagggtttg 1680
ccgtggaaag cggcgaggcg ctggagccga ccctgcgcgc ggcgatggac gtcgacggcc 1740
cggcggtagt ggccatccc gtggattatc gcgataacc gctgctgatg ggccagctgc 1800
atctgagtca gattctgtaa gtcatacaaa taaggaaaga aaaatgaaaa aagtcgact 1860
tgttaccggc gccggccagg ggattggtaa agctatcgcc cttcgtctgg tgaaggatgg 1920
atttgccgtg gccattgcc attataacga cgccaccgcc aaagcggtcg cctccgaaat 1980
caaccaggcc ggcgccgcg ccatggcggg gaaagtggat gtttctgacc gcgaccaggt 2040
atttgccgcc gtcga 2055

<210> 5
<211> 554
<212> PRT
<213> Lactococcus lactis

<400> 5

Met Ser Glu Lys Gln Phe Gly Ala Asn Leu Val Val Asp Ser Leu Ile
1 5 10 15

Asn His Lys Val Lys Tyr Val Phe Gly Ile Pro Gly Ala Lys Ile Asp
20 25 30

Arg Val Phe Asp Leu Leu Glu Asn Glu Glu Gly Pro Gln Met Val Val
35 40 45

Thr Arg His Glu Gln Gly Ala Ala Phe Met Ala Gln Ala Val Gly Arg
50 55 60

Leu Thr Gly Glu Pro Gly Val Val Val Val Thr Ser Gly Pro Gly Val
Page 7

340

345

350

Gln His Glu Phe Asp Thr Glu Asn Thr Glu Glu Gly Arg Met His Pro
 355 360 365

Leu Asp Leu Val Ser Thr Phe Gln Glu Ile Val Lys Asp Asp Glu Thr
 370 375 380

Val Thr Val Asp Val Gly Ser Leu Tyr Ile Trp Met Ala Arg His Phe
 385 390 395 400

Lys Ser Tyr Glu Pro Arg His Leu Leu Phe Ser Asn Gly Met Gln Thr
 405 410 415

Leu Gly Val Ala Leu Pro Trp Ala Ile Thr Ala Ala Leu Leu Arg Pro
 420 425 430

Gly Lys Lys Val Tyr Ser His Ser Gly Asp Gly Gly Phe Leu Phe Thr
 435 440 445

Gly Gln Glu Leu Glu Thr Ala Val Arg Leu Asn Leu Pro Ile Val Gln
 450 455 460

Ile Ile Trp Asn Asp Gly His Tyr Asp Met Val Lys Phe Gln Glu Glu
 465 470 475 480

Met Lys Tyr Gly Arg Ser Ala Ala Val Asp Phe Gly Tyr Val Asp Tyr
 485 490 495

Val Lys Tyr Ala Glu Ala Met Arg Ala Lys Gly Tyr Arg Ala His Ser
 500 505 510

Lys Glu Glu Leu Ala Glu Ile Leu Lys Ser Ile Pro Asp Thr Thr Gly
 515 520 525

Pro Val Val Ile Asp Val Pro Leu Asp Tyr Ser Asp Asn Ile Lys Leu
 530 535 540

Ala Glu Lys Leu Leu Pro Glu Glu Phe Tyr
 545 550

<210> 6
 <211> 3220
 <212> DNA
 <213> Lactococcus lactis

<400> 6
 tagatccgga aacaactgat tacctgagtt aacttagcag aaattgcaga agataacggt 60
 aatttgatg aagcattaaa ttacctttat caaattccgg tgaatgatga aaattatatt 120
 gctgctttaa tcaaaattgc tgacttatat caatttgaag ttgattttga aacagcaatt 180
 tctaagttag aagaagcaag agaattatcg gattctcctc tgattacttt tgctttggct 240

20130927_CL5646WOPCT_ST25.txt

gagtcctact ttgaacaagg tgattattca gctgccatta ccgaatatgc aaaactttca	300
gaacgaaaaa ttttacatga aacaaaaatt tctattttatc aaagaattgg tgactcttat	360
gcccaattag gtaattttga gaatgccata tcatthcttg aaaaatcact tgaatttgat	420
gaaaaaccgg aaaccttgta taaaattgct cttctttatg gagaaactca taatgaaaca	480
agagccattg ctaattttcaa acggtttagaa aaaatggatg ttgaattttt gaactatgaa	540
ttagcctatg cccaaaccct agaagctaata caagaattta aagctgcact agaaatggca	600
aagaaagggga tgaaaaaaaa tcctaatagcc gttcctctct tacacttcgc ttcaaaaatt	660
tgtttcaaac ttaaggacaa agctgcagca gaacgttatc tcgtggatgc tttaaattta	720
ccagaattac atgacgaaac agtctttttg cttgctaatt tatacttcaa cgaagaagat	780
tttgaagctg tcattaatct tgaagagctt ttagaagatg aacattttatt agctaaatgg	840
ctttttgcag gagcacataa agctttggaa aatgattctg aagcggctgc tttgtatgaa	900
gaactcattc aaaccaatct gtcagagaat ccagagtttt tagaagacta tattgatttt	960
cttaaagaaa ttggtcaaat ttctaaaaca gaaccaatta ttgaacaata tttggaactt	1020
gttcagatg atgaaaatat gagaaattta ctgacagact taaaaaataa ttactgacaa	1080
agctgtcagt aattattttt attgtaagct agaaaattca aaaacttgcg tcaaaataat	1140
tgtaaaagggt tctattatct gataaaatga ttgtgaagta atccaagaga ttatgaaata	1200
tgaattagaa caaatagagg taaaataaaa aatgtctgag aaacaatttg gggcgaactt	1260
ggttgtcagt agtttgatta accataaagt gaagtatgta tttgggattc caggagcaaa	1320
aattgaccgg gtttttgatt tattagaaaa tgaagaaggc cctcaaattg tcgtgactcg	1380
tcattgacaa ggagctgctt tcattggtca agctgtcggg cgtttaactg gcgaacctgg	1440
tgtatgattt gttacgagtg ggcctgggtg atcaaaccct gcgactccgc ttttgaccgc	1500
gacatcagaa ggtgatgcta ttttggttat cgggtggacaa gttaaacgaa gtgaccgtct	1560
taaacgtgcg caccaatcaa tggataatgc tggaaatgat caatcagcaa caaaatatct	1620
agcagaagtt cttgacccta atacactttc tgaatcaatt gccaacgctt atcgtattgc	1680
aaaatcagga catccagggt caactttctt atcaatcccc caagatgtaa cggatgccga	1740
agtatcaatc aaagccattc aaccactttc agaccctaaa atggggaatg cctctattga	1800
tgacattaat tatttagcac aagcaattaa aaatgctgta ttgccagtaa ttttggttgg	1860
agctggtgct tcagatgcta aagtcgcttc atccttgctt aatctattga ctcatgttaa	1920
tattcctgtc gttgaaacat tccaagggtg aggggttatt tcacatgatt tagaacatac	1980
tttttatgga cgtatcgggtc ttttccgcaa tcaaccaggc gatatgcttc tgaaacgttc	2040
tgaccttggt attgctgttg gttatgacct aattgaatat gaagctcgta actggaatgc	2100
agaaattgat agtcgaatta tcgttattga taatgccatt gctgaaattg atacttacta	2160
ccaaccagag cgtgaattaa ttggtgatat cgcagcaaca ttggataatc ttttaccagc	2220
tgttcgtggc taaaaattc caaaaggaac aaaagattat ctcgatggcc ttcatgaagt	2280

20130927_CL5646WOPCT_ST25. txt

tgctgagcaa cacgaatttg atactgaaaa tactgaagaa ggtagaatgc accctcttga	2340
tttggtcagc actttccaag aaatcgtcaa ggatgatgaa acagtaaccg ttgacgtagg	2400
ttcactctac atttggatgg cacgtcattt caaatcatat gaaccacgtc atctcctctt	2460
ctcaaacgga atgcaaacac tcggagttgc acttccttgg gcaattacag ccgcattggt	2520
gcgcccaggt aaaaaagttt attcacactc tggatgatgaa ggcttccttt tcacagggca	2580
agaattggaa acagctgtac gtttgaatct tccaatcgtt caaattatct ggaatgacgg	2640
ccattatgat atgggttaaat tccaagaaga aatgaaatat ggctcgttcag cagccgttga	2700
ttttggctat gttgattacg taaaatatgc tgaagcaatg agagcaaaag gttaccgtgc	2760
acacagcaaa gaagaacttg ctgaaattct caaatcaatc ccagatacta ctggaccggt	2820
ggtaattgac gttcctttgg actattctga taacattaaa ttagcagaaa aattattgcc	2880
tgaagagttt tattgattac aatcaagcaa tttgtggcat aacaaaataa aagaagaagg	2940
ccttgaacac ctaagcgttc agggcctttt tttgtgaaat aaattagatg aaatttacia	3000
tgagttttgt gaaactagct tctagtttgt gaaaaattgc ctataattgc cgaataaaaa	3060
taccatttta ccactccaag aggatgcttc aaattagcta aataccgtt ttagaggatg	3120
cgtaaaaaca acaaaagagg atgagtatag aacgataaaa cttttttatg ataggttgag	3180
agaattgaat ataaaatata ataagtagaa ggcagcaatt	3220

<210> 7
 <211> 491
 <212> PRT
 <213> Escheri ch i a col i

<400> 7

Met	Ala	Asn	Tyr	Phe	Asn	Thr	Leu	Asn	Leu	Arg	Gln	Gln	Leu	Ala	Gln
1				5					10					15	

Leu	Gly	Lys	Cys	Arg	Phe	Met	Gly	Arg	Asp	Glu	Phe	Ala	Asp	Gly	Ala
		20						25					30		

Ser	Tyr	Leu	Gln	Gly	Lys	Lys	Val	Val	Ile	Val	Gly	Cys	Gly	Ala	Gln
		35					40					45			

Gly	Leu	Asn	Gln	Gly	Leu	Asn	Met	Arg	Asp	Ser	Gly	Leu	Asp	Ile	Ser
	50					55					60				

Tyr	Ala	Leu	Arg	Lys	Glu	Ala	Ile	Ala	Glu	Lys	Arg	Ala	Ser	Trp	Arg
65					70					75				80	

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

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

20130927_CL5646W0PCT_ST25. txt

Asp Val Val Arg Thr Val Gln Pro Leu Met Lys Asp Gly Ala Ala Leu
115 120 125

Gly Tyr Ser His Gly Phe Asn Ile Val Glu Val Gly Glu Gln Ile Arg
130 135 140

Lys Asp Ile Thr Val Val Met Val Ala Pro Lys Cys Pro Gly Thr Glu
145 150 155 160

Val Arg Glu Glu Tyr Lys Arg Gly Phe Gly Val Pro Thr Leu Ile Ala
165 170 175

Val His Pro Glu Asn Asp Pro Lys Gly Glu Gly Met Ala Ile Ala Lys
180 185 190

Ala Trp Ala Ala Ala Thr Gly Gly His Arg Ala Gly Val Leu Glu Ser
195 200 205

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

Leu Cys Gly Met Leu Gln Ala Gly Ser Leu Leu Cys Phe Asp Lys Leu
225 230 235 240

Val Glu Glu Gly Thr Asp Pro Ala Tyr Ala Glu Lys Leu Ile Gln Phe
245 250 255

Gly Trp Glu Thr Ile Thr Glu Ala Leu Lys Gln Gly Gly Ile Thr Leu
260 265 270

Met Met Asp Arg Leu Ser Asn Pro Ala Lys Leu Arg Ala Tyr Ala Leu
275 280 285

Ser Glu Gln Leu Lys Glu Ile Met Ala Pro Leu Phe Gln Lys His Met
290 295 300

Asp Asp Ile Ile Ser Gly Glu Phe Ser Ser Gly Met Met Ala Asp Trp
305 310 315 320

Ala Asn Asp Asp Lys Lys Leu Leu Thr Trp Arg Glu Glu Thr Gly Lys
325 330 335

Thr Ala Phe Glu Thr Ala Pro Gln Tyr Glu Gly Lys Ile Gly Glu Gln
340 345 350

Glu Tyr Phe Asp Lys Gly Val Leu Met Ile Ala Met Val Lys Ala Gly
355 360 365

Val Glu Leu Ala Phe Glu Thr Met Val Asp Ser Gly Ile Ile Glu Glu
370 375 380

20130927_CL5646WOPCT_ST25. txt

Ser Ala Tyr Tyr Glu Ser Leu His Glu Leu Pro Leu Ile Ala Asn Thr
385 390 395 400

Ile Ala Arg Lys Arg Leu Tyr Glu Met Asn Val Val Ile Ser Asp Thr
405 410 415

Ala Glu Tyr Gly Asn Tyr Leu Phe Ser Tyr Ala Cys Val Pro Leu Leu
420 425 430

Lys Pro Phe Met Ala Glu Leu Gln Pro Gly Asp Leu Gly Lys Ala Ile
435 440 445

Pro Glu Gly Ala Val Asp Asn Gly Gln Leu Arg Asp Val Asn Glu Ala
450 455 460

Ile Arg Ser His Ala Ile Glu Gln Val Gly Lys Lys Leu Arg Gly Tyr
465 470 475 480

Met Thr Asp Met Lys Arg Ile Ala Val Ala Gly
485 490

<210> 8
<211> 1476
<212> DNA
<213> Escherichia coli

<400> 8
atggctaact acttcaatac actgaatctg cgccagcagc tggcacagct gggcaaatgt 60
cgctttatgg gccgcgatga attcgccgat ggcgcgagct accttcaggg taaaaaagta 120
gtcatcgtcg gctgtggcgc acagggctctg aaccagggcc tgaacatgcg tgattctggt 180
ctcgatatct cctacgctct gcgtaaagaa gcgattgccg agaagcgcgc gtcctggcgt 240
aaagcgaccg aaaatggttt taaagtgggt acttacgaag aactgatccc acaggcggat 300
ctggtgatta acctgacgcc ggacaagcag cactctgatg tagtgcgcac cgtacagcca 360
ctgatgaaag acggcgcggc gctgggctac tcgcacgggt tcaacatcgt cgaagtgggc 420
gagcagatcc gtaaagatat caccgtagtg atggttgccg cgaaatgccc aggaccgaa 480
gtgcgtgaag agtacaacg tgggttcggc gtaccgacgc tgattgccgt tcacccgaa 540
aacgatccga aaggcgaagg catggcgatt gccaaagcct gggcggctgc aaccggtggt 600
caccgtgcgg gtgtgctgga atcgtccttc gttgcggaag tgaaatctga cctgatgggc 660
gagcaaacca tcctgtgcgg tatgttgagc gctggctctc tgctgtgctt cgacaagctg 720
gtggaagaag gtaccgatcc agcatagca gaaaaactga ttcagttcgg ttgggaaacc 780
atcaccgaag cactgaaaca gggcggcatc accctgatga tggaccgtct ctctaaccgc 840
gcgaaactgc gtgcttatgc gctttctgaa cagctgaaag agatcatggc acccctgttc 900
cagaaacata tggacgacat catctccggc gaattctctt ccggtatgat ggcggactgg 960
gccaacgatg ataagaaact gctgacctgg cgtgaagaga cgggcaaac cgcgtttgaa 1020

20130927_CL5646WOPCT_ST25.txt

accgcgccgc agtatgaagg caaaatcggc gagcaggagt acttcgataa aggcgtactg 1080
atgattgcga tgggtgaaagc gggcgttgaa ctggcgttcg aaaccatggt cgattccggc 1140
atcattgaag agtctgcata ttatgaatca ctgcacgagc tgccgctgat tgccaacacc 1200
atcgcccgtg agcgtctgta cgaaatgaac gtggttatct ctgataccgc tgagtacggt 1260
aactatctgt tctcttacgc ttgtgtgccg ttgctgaaac cgtttatggc agagctgcaa 1320
ccgggcgacc tgggtaaagc tattccggaa ggcgcggtag ataacgggca actgctgat 1380
gtgaacgaag cgattcgag ccatgcgatt gagcaggtag gtaagaaact gcgcggctat 1440
atgacagata tgaaacgtat tgctgttcg gggttaa 1476

<210> 9
<211> 395
<212> PRT
<213> Saccharomyces cerevisiae

<400> 9

Met Leu Arg Thr Gln Ala Ala Arg Leu Ile Cys Asn Ser Arg Val Ile
1 5 10 15

Thr Ala Lys Arg Thr Phe Ala Leu Ala Thr Arg Ala Ala Ala Tyr Ser
20 25 30

Arg Pro Ala Ala Arg Phe Val Lys Pro Met Ile Thr Thr Arg Gly Leu
35 40 45

Lys Gln Ile Asn Phe Gly Gly Thr Val Glu Thr Val Tyr Glu Arg Ala
50 55 60

Asp Trp Pro Arg Glu Lys Leu Leu Asp Tyr Phe Lys Asn Asp Thr Phe
65 70 75 80

Ala Leu Ile Gly Tyr Gly Ser Gln Gly Tyr Gly Gln Gly Leu Asn Leu
85 90 95

Arg Asp Asn Gly Leu Asn Val Ile Ile Gly Val Arg Lys Asp Gly Ala
100 105 110

Ser Trp Lys Ala Ala Ile Glu Asp Gly Trp Val Pro Gly Lys Asn Leu
115 120 125

Phe Thr Val Glu Asp Ala Ile Lys Arg Gly Ser Tyr Val Met Asn Leu
130 135 140

Leu Ser Asp Ala Ala Gln Ser Glu Thr Trp Pro Ala Ile Lys Pro Leu
145 150 155 160

Leu Thr Lys Gly Lys Thr Leu Tyr Phe Ser His Gly Phe Ser Pro Val
165 170 175

20130927_CL5646W0PCT_ST25. txt

Phe Lys Asp Leu Thr His Val Glu Pro Pro Lys Asp Leu Asp Val Ile
180 185 190

Leu Val Ala Pro Lys Gly Ser Gly Arg Thr Val Arg Ser Leu Phe Lys
195 200 205

Glu Gly Arg Gly Ile Asn Ser Ser Tyr Ala Val Trp Asn Asp Val Thr
210 215 220

Gly Lys Ala His Glu Lys Ala Gln Ala Leu Ala Val Ala Ile Gly Ser
225 230 235 240

Gly Tyr Val Tyr Gln Thr Thr Phe Glu Arg Glu Val Asn Ser Asp Leu
245 250 255

Tyr Gly Glu Arg Gly Cys Leu Met Gly Gly Ile His Gly Met Phe Leu
260 265 270

Ala Gln Tyr Asp Val Leu Arg Glu Asn Gly His Ser Pro Ser Glu Ala
275 280 285

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

Gly Lys Tyr Gly Met Asp Tyr Met Tyr Asp Ala Cys Ser Thr Thr Ala
305 310 315 320

Arg Arg Gly Ala Leu Asp Trp Tyr Pro Ile Phe Lys Asn Ala Leu Lys
325 330 335

Pro Val Phe Gln Asp Leu Tyr Glu Ser Thr Lys Asn Gly Thr Glu Thr
340 345 350

Lys Arg Ser Leu Glu Phe Asn Ser Gln Pro Asp Tyr Arg Glu Lys Leu
355 360 365

Glu Lys Glu Leu Asp Thr Ile Arg Asn Met Glu Ile Trp Lys Val Gly
370 375 380

Lys Glu Val Arg Lys Leu Arg Pro Glu Asn Gln
385 390 395

<210> 10
<211> 1188
<212> DNA
<213> Saccharomyces cerevisiae

<400> 10
atgttgagaa ctcaagccgc cagattgatc tgcaactccc gtgtcatcac tgctaagaga 60
acctttgctt tggccaccgc tgctgctgct tacagcagac cagctgcccg tttcgttaag 120
ccaatgatca ctaccgctgg tttgaagcaa atcaacttcg gtggtactgt tgaaaccgctc 180

```

tacgaaagag ctgactggcc aagagaaaag ttgttggact acttcaagaa cgacactttt      240
gctttgatcg gttacggttc ccaaggttac ggtcaagggt tgaacttgag agacaacggt      300
ttgaacgtta tcattggtgt ccgtaaagat ggtgcttctt ggaaggctgc catcgaagac      360
ggttgggttc caggcaagaa cttgttctact gttgaagatg ctatcaagag aggtagttac      420
gttatgaact tgttgtccga tgccgctcaa tcagaaacct ggcttgctat caagccattg      480
ttgaccaagg gtaagacttt gtacttctcc cacggtttct cccagttctt caaggacttg      540
actcacgttg aaccaccaaa ggacttagat gttatcttgg ttgctccaaa gggttccggt      600
agaactgtca gatctttgtt caaggaaggt cgtggtatta actcttctta cgccgtctgg      660
aacgatgtca ccggttaaggc tcacgaaaag gcccaagctt tggccggttg cattggttcc      720
ggttacgttt accaaaccac tttcgaaaga gaagtcaact ctgacttgta cggtgaaaga      780
ggttgtttaa tgggtggtat ccacggtatg ttcttggctc aatacgacgt cttgagagaa      840
aacggtcact ccccatctga agctttcaac gaaaccgtcg aagaagctac ccaatctcta      900
taccattga tcggttaagta cggtatggat tacatgtacg atgcttggtc caccaccgcc      960
agaagaggtg ctttggactg gtaccaatc ttcaagaatg ctttgaagcc tgttttccaa     1020
gacttgtacg aatctaccaa gaacggtacc gaaaccaaga gatctttgga attcaactct     1080
caacctgact acagagaaaa gctagaaaag gaattagaca ccatcagaaa catggaaatc     1140
tggaagggtg gtaaggaagt cagaaagttg agaccagaaa accaataa                     1188

```

```

<210> 11
<211> 330
<212> PRT
<213> Methanococcus mari pal udi s

```

```
<400> 11
```

```
Met Lys Val Phe Tyr Asp Ser Asp Phe Lys Leu Asp Al a Leu Lys Gl u
1          5          10          15
```

```
Lys Thr Ile Al a Val Ile Gly Tyr Gly Ser Gl n Gly Arg Al a Gl n Ser
          20          25          30
```

```
Leu Asn Met Lys Asp Ser Gly Leu Asn Val Val Val Gly Leu Arg Lys
          35          40          45
```

```
Asn Gly Al a Ser Trp Asn Asn Al a Lys Al a Asp Gly Hi s Asn Val Met
          50          55          60
```

```
Thr Ile Gl u Gl u Al a Al a Gl u Lys Al a Asp Ile Ile Hi s Ile Leu Ile
65          70          75          80
```

```
Pro Asp Gl u Leu Gl n Al a Gl u Val Tyr Gl u Ser Gl n Ile Lys Pro Tyr
          85          90          95
```

```
Leu Lys Gl u Gly Lys Thr Leu Ser Phe Ser Hi s Gly Phe Asn Ile Hi s
```

100

105

110

Tyr Gly Phe Ile Val Pro Pro Lys Gly Val Asn Val Val Leu Val Ala
 115 120 125

Pro Lys Ser Pro Gly Lys Met Val Arg Arg Thr Tyr Glu Glu Gly Phe
 130 135 140

Gly Val Pro Gly Leu Ile Cys Ile Glu Ile Asp Ala Thr Asn Asn Ala
 145 150 155 160

Phe Asp Ile Val Ser Ala Met Ala Lys Gly Ile Gly Leu Ser Arg Ala
 165 170 175

Gly Val Ile Gln Thr Thr Phe Lys Glu Glu Thr Glu Thr Asp Leu Phe
 180 185 190

Gly Glu Gln Ala Val Leu Cys Gly Gly Val Thr Glu Leu Ile Lys Ala
 195 200 205

Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Ala Pro Glu Met Ala Tyr
 210 215 220

Phe Glu Thr Cys His Glu Leu Lys Leu Ile Val Asp Leu Ile Tyr Gln
 225 230 235 240

Lys Gly Phe Lys Asn Met Trp Asn Asp Val Ser Asn Thr Ala Glu Tyr
 245 250 255

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

Ala Met Lys Glu Ile Leu Arg Glu Ile Gln Asp Gly Arg Phe Thr Lys
 275 280 285

Glu Phe Leu Leu Glu Lys Gln Val Ser Tyr Ala His Leu Lys Ser Met
 290 295 300

Arg Arg Leu Glu Gly Asp Leu Gln Ile Glu Glu Val Gly Ala Lys Leu
 305 310 315 320

Arg Lys Met Cys Gly Leu Glu Lys Glu Glu
 325 330

<210> 12
 <211> 993
 <212> DNA
 <213> Methanococcus maripaludis

<400> 12
 atgaaggat tctatgactc agattttaaa ttagatgctt taaaagaaaa aacaattgca 60

gtaatcgggt atggaagtca aggtagggca cagtccttaa acatgaaaga cagcggatta 120

```

aacgttggtg ttggtttaag aaaaaacggt gcttcatgga acaacgctaa agcagacggt      180
cacaatgtaa tgaccattga agaagctgct gaaaaagcgg acatcatcca catcttaata      240
cctgatgaat tacaggcaga agtttatgaa agccagataa aaccatacct aaaagaagga      300
aaaacactaa gcttttcaca tggttttaac atccactatg gattcattgt tccacaaaaa      360
ggagttaacg tggtttttagt tgctccaaaa tcacctggaa aaatggttag aagaacatac      420
gaagaagggt tcggtgttcc aggtttaatc tgtattgaaa ttgatgcaac aaacaacgca      480
tttgatattg tttcagcaat ggcaaaagga atcggtttat caagagctgg agttatccag      540
acaactttca aagaagaaac agaaactgac cttttcggtg aacaagctgt tttatgcggt      600
ggagttaccg aattaatcaa ggcaggattt gaaacactcg ttgaagcagg atacgcacca      660
gaaatggcat actttgaaac ctgccacgaa ttgaaattaa tcgttgactt aatctaccaa      720
aaaggattca aaaacatgtg gaacgatgta agtaacactg cagaatacgg cggacttaca      780
agaagaagca gaatcgttac agctgattca aaagctgcaa tgaaagaaat ctaagagaa      840
atccaagatg gaagattcac aaaagaattc cttctcgaaa aacaggtaag ctatgctcat      900
ttaaatacaa tgagaagact cgaaggagac ttacaaatcg aagaagtcgg cgcaaaatta      960
agaaaaatgt gcggtcttga aaaagaagaa taa                                  993

```

```

<210> 13
<211> 342
<212> PRT
<213> Bacillus subtilis

```

```
<400> 13
```

```
Met Val Lys Val Tyr Tyr Asn Gly Asp Ile Lys Glu Asn Val Leu Ala
1          5          10          15
```

```
Gly Lys Thr Val Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His
20          25          30
```

```
Ala Leu Asn Leu Lys Glu Ser Gly Val Asp Val Ile Val Gly Val Arg
35          40          45
```

```
Gln Gly Lys Ser Phe Thr Gln Ala Gln Glu Asp Gly His Lys Val Phe
50          55          60
```

```
Ser Val Lys Glu Ala Ala Ala Gln Ala Glu Ile Ile Met Val Leu Leu
65          70          75          80
```

```
Pro Asp Glu Gln Gln Gln Lys Val Tyr Glu Ala Glu Ile Lys Asp Glu
85          90          95
```

```
Leu Thr Ala Gly Lys Ser Leu Val Phe Ala His Gly Phe Asn Val His
100          105          110
```

```
Phe His Gln Ile Val Pro Pro Ala Asp Val Asp Val Phe Leu Val Ala
Page 18

```

115

120

125

Pro Lys Gly Pro Gly His Leu Val Arg Arg Thr Tyr Glu Gln Gly Ala
 130 135 140

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

Arg Asp Lys Ala Leu Ala Tyr Ala Lys Gly Ile Gly Gly Ala Arg Ala
 165 170 175

Gly Val Leu Glu Thr Thr Phe Lys Glu Glu Thr Glu Thr Asp Leu Phe
 180 185 190

Gly Glu Gln Ala Val Leu Cys Gly Gly Leu Ser Ala Leu Val Lys Ala
 195 200 205

Gly Phe Glu Thr Leu Thr Glu Ala Gly Tyr Gln Pro Glu Leu Ala Tyr
 210 215 220

Phe Glu Cys Leu His Glu Leu Lys Leu Ile Val Asp Leu Met Tyr Glu
 225 230 235 240

Glu Gly Leu Ala Gly Met Arg Tyr Ser Ile Ser Asp Thr Ala Gln Trp
 245 250 255

Gly Asp Phe Val Ser Gly Pro Arg Val Val Asp Ala Lys Val Lys Glu
 260 265 270

Ser Met Lys Glu Val Leu Lys Asp Ile Gln Asn Gly Thr Phe Ala Lys
 275 280 285

Glu Trp Ile Val Glu Asn Gln Val Asn Arg Pro Arg Phe Asn Ala Ile
 290 295 300

Asn Ala Ser Glu Asn Glu His Gln Ile Glu Val Val Gly Arg Lys Leu
 305 310 315 320

Arg Glu Met Met Pro Phe Val Lys Gln Gly Lys Lys Lys Glu Ala Val
 325 330 335

Val Ser Val Ala Gln Asn
 340

<210> 14
 <211> 1476
 <212> DNA
 <213> Bacillus subtilis

<400> 14
 atggctaact acttcaatac actgaatctg cgccagcagc tggcacagct gggcaaagt 60

cgctttatgg gccgcgatga attcgccgat ggcgcgagct accttcaggg taaaaaagta 120

20130927_CL5646WOPCT_ST25.txt

```

gtcatcgtcg gctgtggcgc acagggctctg aaccagggcc tgaacatgcg tgattctggt      180
ctcgatatct cctacgctct gcgtaaagaa gcgattgccg agaagcgcgcg gtcctggcgt      240
aaagcgaccg aaaatggttt taaagtgggt acttacgaag aactgatccc acaggcggat      300
ctggtgatta acctgacgcc ggacaagcag cactctgatg tagtgcgcac cgtacagcca      360
ctgatgaaag acggcgcggc gctgggctac tcgcacgggt tcaacatcgt cgaagtgggc      420
gagcagatcc gtaaagatat caccgtagtg atggttgcgc cgaaatgccc aggcaccgaa      480
gtgctgaag agtacaaacg tgggttcggc gtaccgacgc tgattgccgt tcacccggaa      540
aacgatccga aaggcgaagg catggcgatt gccaaagcct gggcggctgc aaccggtggt      600
caccgtgcgg gtgtgctgga atcgctcttc gttgcggaag tgaaatctga cctgatgggc      660
gagcaaacca tcctgtgcgg tatgttgagc gctggctctc tgctgtgctt cgacaagctg      720
gtggaagaag gtaccgatcc agcatacgca gaaaaactga ttcagttcgg ttgggaaacc      780
atcaccgaag cactgaaaca gggcggcatc accctgatga tggaccgtct ctctaaccgc      840
gcgaaactgc gtgcttatgc gctttctgaa cagctgaaag agatcatggc acccctgttc      900
cagaaacata tggacgacat catctccggc gaattctctt ccggtatgat ggcggactgg      960
gccaacgatg ataagaaact gctgacctgg cgtgaagaga ccggcaaaac cgcgtttgaa     1020
accgcgccgc agtatgaagg caaaatcggc gagcaggagt acttcgataa aggcgtactg     1080
atgattgcga tggtgaaagc gggcgttgaa ctggcgttcg aaaccatggt cgattccggc     1140
atcattgaag agtctgcata ttatgaatca ctgcacgagc tgccgctgat tgccaacacc     1200
atcgcccgtg agcgtctgta cgaaatgaac gtggttatct ctgataccgc tgagtacggt     1260
aactatctgt tctcttacgc ttgtgtgccg ttgctgaaac cgtttatggc agagctgcaa     1320
ccgggcgacc tgggtaaagc tattccggaa ggcgcggtag ataacgggca actgcgtgat     1380
gtgaacgaag cgattcgag ccatgcgatt gagcaggtag gtaagaaact gcgcggctat     1440
atgacagata tgaaacgtat tgctgttgcg gggttaa                                1476

```

<210> 15
 <211> 343
 <212> PRT
 <213> Anaerostipes caccae

<400> 15

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
 1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
 20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
 35 40 45

Ile Ile Gly Leu Tyr Glu Gly Ala Lys Glu Trp Lys Arg Ala Glu Glu
 Page 20

50

55

60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

Ile Ile Met Ile Leu Ile Asn Asp Glu Lys Gln Ala Thr Met Tyr Lys
85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
210 215 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225 230 235 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
245 250 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
260 265 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
275 280 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
Page 21

325

330

335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 16
<211> 343
<212> PRT
<213> Anaerostipes caccae

<400> 16

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Tyr Glu Gly Ala Lys Asp Trp Lys Arg Ala Glu Glu
50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

Ile Ile Met Ile Leu Ile Asn Asp Glu Lys Gln Ala Thr Met Tyr Lys
85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
210 215 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225 230 235 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
245 250 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
260 265 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
275 280 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 17
<211> 616
<212> PRT
<213> Escherichia coli

<400> 17

Met Pro Lys Tyr Arg Ser Ala Thr Thr Thr His Gly Arg Asn Met Ala
1 5 10 15

Gly Ala Arg Ala Leu Trp Arg Ala Thr Gly Met Thr Asp Ala Asp Phe
20 25 30

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

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

Glu Ala Ala Gly Gly Val Ala Lys Glu Phe Asn Thr Ile Ala Val Asp
65 70 75 80

Asp Gly Ile Ala Met Gly His Gly Gly Met Leu Tyr Ser Leu Pro Ser
85 90 95

Arg Glu Leu Ile Ala Asp Ser Val Glu Tyr Met Val Asn Ala His Cys
100 105 110

20130927_CL5646W0PCT_ST25. txt

Ala Asp Ala Met Val Cys Ile Ser Asn Cys Asp Lys Ile Thr Pro Gly
115 120 125

Met Leu Met Ala Ser Leu Arg Leu Asn Ile Pro Val Ile Phe Val Ser
130 135 140

Gly Gly Pro Met Glu Ala Gly Lys Thr Lys Leu Ser Asp Gln Ile Ile
145 150 155 160

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

Ser Asp Ser Gln Ser Asp Gln Val Glu Arg Ser Ala Cys Pro Thr Cys
180 185 190

Gly Ser Cys Ser Gly Met Phe Thr Ala Asn Ser Met Asn Cys Leu Thr
195 200 205

Glu Ala Leu Gly Leu Ser Gln Pro Gly Asn Gly Ser Leu Leu Ala Thr
210 215 220

His Ala Asp Arg Lys Gln Leu Phe Leu Asn Ala Gly Lys Arg Ile Val
225 230 235 240

Glu Leu Thr Lys Arg Tyr Tyr Glu Gln Asn Asp Glu Ser Ala Leu Pro
245 250 255

Arg Asn Ile Ala Ser Lys Ala Ala Phe Glu Asn Ala Met Thr Leu Asp
260 265 270

Ile Ala Met Gly Gly Ser Thr Asn Thr Val Leu His Leu Leu Ala Ala
275 280 285

Ala Gln Glu Ala Glu Ile Asp Phe Thr Met Ser Asp Ile Asp Lys Leu
290 295 300

Ser Arg Lys Val Pro Gln Leu Cys Lys Val Ala Pro Ser Thr Gln Lys
305 310 315 320

Tyr His Met Glu Asp Val His Arg Ala Gly Gly Val Ile Gly Ile Leu
325 330 335

Gly Glu Leu Asp Arg Ala Gly Leu Leu Asn Arg Asp Val Lys Asn Val
340 345 350

Leu Gly Leu Thr Leu Pro Gln Thr Leu Glu Gln Tyr Asp Val Met Leu
355 360 365

Thr Gln Asp Asp Ala Val Lys Asn Met Phe Arg Ala Gly Pro Ala Gly
370 375 380

20130927_CL5646W0PCT_ST25. txt

I l e Arg Thr Thr G l n A l a P h e Ser G l n Asp C y s Arg Trp Asp Thr Leu
385 390 395 400

Asp Asp Asp Arg A l a A s n G l y C y s I l e Arg Ser Leu G l u H i s A l a Tyr
405 410 415

Ser Lys Asp G l y G l y Leu A l a Val Leu Tyr G l y A s n P h e A l a G l u A s n
420 425 430

G l y C y s I l e Val Lys Thr A l a G l y Val Asp Asp Ser I l e Leu Lys P h e
435 440 445

Thr G l y Pro A l a Lys Val Tyr G l u Ser G l n Asp Asp A l a Val G l u A l a
450 455 460

I l e Leu G l y G l y Lys Val Val A l a G l y Asp Val Val Val I l e Arg Tyr
465 470 475 480

G l u G l y Pro Lys G l y G l y Pro G l y Met G l n G l u Met Leu Tyr Pro Thr
485 490 495

Ser P h e Leu Lys Ser Met G l y Leu G l y Lys A l a C y s A l a Leu I l e Thr
500 505 510

Asp G l y Arg P h e Ser G l y G l y Thr Ser G l y Leu Ser I l e G l y H i s Val
515 520 525

Ser Pro G l u A l a A l a Ser G l y G l y Ser I l e G l y Leu I l e G l u Asp G l y
530 535 540

Asp Leu I l e A l a I l e Asp I l e Pro A s n Arg G l y I l e G l n Leu G l n Val
545 550 555 560

Ser Asp A l a G l u Leu A l a A l a Arg Arg G l u A l a G l n Asp A l a Arg G l y
565 570 575

Asp Lys A l a Trp Thr Pro Lys A s n Arg G l u Arg G l n Val Ser P h e A l a
580 585 590

Leu Arg A l a Tyr A l a Ser Leu A l a Thr Ser A l a Asp Lys G l y A l a Val
595 600 605

Arg Asp Lys Ser Lys Leu G l y G l y
610 615

<210> 18

<211> 1851

<212> DNA

<213> Escheri ch i a col i

<400> 18

atgcctaagt accgttccgc caccaccact catggtcgta atatggcggg tgctcgtgcg

60

20130927_CL5646WOPCT_ST25. txt

```

ctgtggcgcg ccaccggaat gaccgacgcc gatttcggta agccgattat cgcggttggt 120
aactcgttca cccaatttgt accgggtcac gtccatctgc gcatctcgg taaactggtc 180
gccgaacaaa ttgaagcggc tggcggcggt gccaaagagt tcaacaccat tgcggtggat 240
gatgggattg ccatgggcca cggggggatg ctttattcac tgccatctcg cgaactgata 300
gctgattccg ttgagtatat ggtcaacgcc cactgcgccg acgcatggt ctgcatctct 360
aactgcgaca aaatcacccc ggggatgctg atggcttccc tgcgcctgaa tattccggtg 420
atctttgttt ccggcggccc gatggaggcc gggaaaacca aactttccga tcagatcatc 480
aagctcgatc tggttgatgc gatgatccag ggcgcagacc cgaaagtatc tgactccag 540
agcgatcagg ttgaacgttc cgcgtgtccg acctgcggtt cctgctccgg gatgtttacc 600
gctaactcaa tgaactgcct gaccgaagcg ctgggcctgt cgagccggg caacggctcg 660
ctgctggcaa cccacgccga ccgtaagcag ctgttcctta atgctggtaa acgattgtt 720
gaattgacca aacgttatta cgagcaaac gacgaaagt cactgccgcg taatatcgcc 780
agtaaggcgg cgtttgaaaa cgccatgacg ctggatatcg cgatgggtgg atcgactaac 840
accgtacttc acctgctggc ggcggcgag gaagcggaaa tcgacttcac catgagtgat 900
atcgataagc tttcccgcaa ggttccacag ctgtgtaaag ttgcgccgag caccagaaa 960
taccatatgg aagatgttca ccgtgctggt ggtgttatcg gtattctcgg cgaactggat 1020
cgcgcggggt tactgaaccg tgatgtgaaa aacgtacttg gcctgacgtt gccgcaaacg 1080
ctggaacaat acgacgttat gctgaccag gatgacgcgg taaaaaatat gttccgcgca 1140
ggtcctgcag gcattcgtac cacacaggca ttctcgaag attgccgttg ggatacgtg 1200
gacgacgatc gcgccaatgg ctgtatccgc tcgctggaac acgcctacag caaagacggc 1260
ggcctggcgg tgctctacgg taactttgcg gaaaacggct gcatcgtgaa aacggcaggc 1320
gtcgatgaca gcatcctcaa attcaccggc ccggcgaaag tgtacgaaag ccaggacgat 1380
gcggtagaag cgattctcgg cggtaaagtt gtcgccggag atgtggtagt aattcgctat 1440
gaaggcccga aaggcggctc ggggatgcag gaaatgctct acccaaccag cttcctgaaa 1500
tcaatgggtc tcggcaaagc ctgtgcgctg atcaccgacg gtcgtttctc tgggtggcacc 1560
tctggtcttt ccatcgcca cgtctcaccg gaagcggcaa gcggcgagc cattggcctg 1620
attgaagatg gtgacctgat cgctatcgac atcccgaacc gtggcattca gttacaggta 1680
agcgatgccg aactggcggc gcgtcgtgaa gcgcaggacg ctcgaggtga caaagcctgg 1740
acgccgaaaa atcgtgaacg tcaggtctcc ttgcccctgc gtgcttatgc cagcctggca 1800
accagcgccg acaaaggcgc ggtgcgcgat aaatcgaaac tgggggggta a 1851

```

```

<210> 19
<211> 585
<212> PRT
<213> Saccharomyces cerevisiae

```

```
<400> 19
```

20130927_CL5646WOPCT_ST25. txt

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

Arg Cys Val Ala Lys Lys Leu Asn Lys Tyr Ser Tyr Ile Ile Thr Glu
20 25 30

Pro Lys Gly Gln Gly Ala Ser Gln Ala Met Leu Tyr Ala Thr Gly Phe
35 40 45

Lys Lys Glu Asp Phe Lys Lys Pro Gln Val Gly Val Gly Ser Cys Trp
50 55 60

Trp Ser Gly Asn Pro Cys Asn Met His Leu Leu Asp Leu Asn Asn Arg
65 70 75 80

Cys Ser Gln Ser Ile Glu Lys Ala Gly Leu Lys Ala Met Gln Phe Asn
85 90 95

Thr Ile Gly Val Ser Asp Gly Ile Ser Met Gly Thr Lys Gly Met Arg
100 105 110

Tyr Ser Leu Gln Ser Arg Glu Ile Ile Ala Asp Ser Phe Glu Thr Ile
115 120 125

Met Met Ala Gln His Tyr Asp Ala Asn Ile Ala Ile Pro Ser Cys Asp
130 135 140

Lys Asn Met Pro Gly Val Met Met Ala Met Gly Arg His Asn Arg Pro
145 150 155 160

Ser Ile Met Val Tyr Gly Gly Thr Ile Leu Pro Gly His Pro Thr Cys
165 170 175

Gly Ser Ser Lys Ile Ser Lys Asn Ile Asp Ile Val Ser Ala Phe Gln
180 185 190

Ser Tyr Gly Glu Tyr Ile Ser Lys Gln Phe Thr Glu Glu Glu Arg Glu
195 200 205

Asp Val Val Glu His Ala Cys Pro Gly Pro Gly Ser Cys Gly Gly Met
210 215 220

Tyr Thr Ala Asn Thr Met Ala Ser Ala Ala Glu Val Leu Gly Leu Thr
225 230 235 240

Ile Pro Asn Ser Ser Ser Phe Pro Ala Val Ser Lys Glu Lys Leu Ala
245 250 255

Glu Cys Asp Asn Ile Gly Glu Tyr Ile Lys Lys Thr Met Glu Leu Gly
260 265 270

I l e L e u P r o A r g A s p I l e L e u T h r L y s G l u A l a P h e G l u A s n A l a I l e
 275 280 285
 T h r T y r V a l V a l A l a T h r G l y G l y S e r T h r A s n A l a V a l L e u H i s L e u
 290 295 300
 V a l A l a V a l A l a H i s S e r A l a G l y V a l L y s L e u S e r P r o A s p A s p P h e
 305 310 315 320
 G l n A r g I l e S e r A s p T h r T h r P r o L e u I l e G l y A s p P h e L y s P r o S e r
 325 330 335
 G l y L y s T y r V a l M e t A l a A s p L e u I l e A s n V a l G l y G l y T h r G l n S e r
 340 345 350
 V a l I l e L y s T y r L e u T y r G l u A s n A s n M e t L e u H i s G l y A s n T h r M e t
 355 360 365
 T h r V a l T h r G l y A s p T h r L e u A l a G l u A r g A l a L y s L y s A l a P r o S e r
 370 375 380
 L e u P r o G l u G l y G l n G l u I l e I l e L y s P r o L e u S e r H i s P r o I l e L y s
 385 390 395 400
 A l a A s n G l y H i s L e u G l n I l e L e u T y r G l y S e r L e u A l a P r o G l y G l y
 405 410 415
 A l a V a l G l y L y s I l e T h r G l y L y s G l u G l y T h r T y r P h e L y s G l y A r g
 420 425 430
 A l a A r g V a l P h e G l u G l u G l u G l y A l a P h e I l e G l u A l a L e u G l u A r g
 435 440 445
 G l y G l u I l e L y s L y s G l y G l u L y s T h r V a l V a l V a l I l e A r g T y r G l u
 450 455 460
 G l y P r o A r g G l y A l a P r o G l y M e t P r o G l u M e t L e u L y s P r o S e r S e r
 465 470 475 480
 A l a L e u M e t G l y T y r G l y L e u G l y L y s A s p V a l A l a L e u L e u T h r A s p
 485 490 495
 G l y A r g P h e S e r G l y G l y S e r H i s G l y P h e L e u I l e G l y H i s I l e V a l
 500 505 510
 P r o G l u A l a A l a G l u G l y G l y P r o I l e G l y L e u V a l A r g A s p G l y A s p
 515 520 525
 G l u I l e I l e I l e A s p A l a A s p A s n A s n L y s I l e A s p L e u L e u V a l S e r
 530 535 540

20130927_CL5646WOPCT_ST25. txt

Asp Lys Glu Met Ala Gln Arg Lys Gln Ser Trp Val Ala Pro Pro Pro
545 550 555 560

Arg Tyr Thr Arg Gly Thr Leu Ser Lys Tyr Ala Lys Leu Val Ser Asn
565 570 575

Ala Ser Asn Gly Cys Val Leu Asp Ala
580 585

<210> 20
<211> 1131
<212> DNA
<213> Saccharomyces cerevisiae

<400> 20
atgaccttgg caccctaga cgctccaaa gttaagataa ctaccacaca acatgcatct 60
aagccaaaac cgaacagtga gttagtgttt ggcaagagct tcacggacca catgtttaact 120
gcggaatgga cagctgaaaa aggggtgggggt accccagaga ttaaacctta tcaaaatctg 180
tctttagacc cttccgcggt ggttttccat tatgcttttg agctattcga agggatgaag 240
gcttacagaa cgggtggacaa caaaattaca atgtttcgtc cagatatgaa tatgaagcgc 300
atgaataagt ctgctcagag aatctgtttg ccaacgttcg acccagaaga gttgattacc 360
ctaattggga aactgatcca gcaagataag tgcttagttc ctgaaggaaa aggttactct 420
ttatataatca ggcctacatt aatcggcact acggccggtt taggggtttc cagcctgat 480
agagccttgc tatatgtcat ttgctgccct gtgggtcctt attacaaaac tggatttaag 540
gcggtcagac tggaaagccac tgattatgcc acaagagctt ggccaggagg ctgtggtgac 600
aagaaactag gtgcaaaact cgccccctgc gtccctgccac aattgcaagc tgcttcaagg 660
ggttaccac aaattttatg gctatttgggt ccaaataaca acattactga agtcggcacc 720
atgaatgctt ttttcgtgtt taaagatagt aaaacgggca agaaggaact agttactgct 780
ccactagacg gtaccatttt ggaaggtgtt actagggatt ccattttaaa tcttgctaaa 840
gaaagactcg aaccaagtga atggaccatt agtgaacgct acttcactat aggcgaagtt 900
actgagagat ccaagaacgg tgaactactt gaagcctttg gttctggtac tgctgcgatt 960
gtttctcca ttaaggaaat cggctggaaa ggcaacaaa ttaatatcc gttgttgccc 1020
ggcgaacaaa ccggtccatt ggccaaagaa gttgcacaat ggattaatgg aatccaatat 1080
ggcgagactg agcatggcaa ttggtcaagg gttgttactg atttgaactg a 1131

<210> 21
<211> 550
<212> PRT
<213> Methanococcus maripaludis

<400> 21

Met Ile Ser Asp Asn Val Lys Lys Gly Val Ile Arg Thr Pro Asn Arg
1 5 10 15

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

20130927_CL5646WOPCT_ST25.txt

Asp Asp Phe Asp Arg Leu Ser Asp Glu Val Pro His Ile Ala Ser Ile
 290 295 300
 Lys Pro Gly Gly Glu His Tyr Met Ile Asp Leu His Asn Ala Gly Gly
 305 310 315 320
 Ile Pro Ala Val Leu Asn Val Leu Lys Glu Lys Ile Arg Asp Thr Lys
 325 330 335
 Thr Val Asp Gly Arg Ser Ile Leu Glu Ile Ala Glu Ser Val Lys Tyr
 340 345 350
 Ile Asn Tyr Asp Val Ile Arg Lys Val Glu Ala Pro Val His Glu Thr
 355 360 365
 Ala Gly Leu Arg Val Leu Lys Gly Asn Leu Ala Pro Asn Gly Cys Val
 370 375 380
 Val Lys Ile Gly Ala Val His Pro Lys Met Tyr Lys His Asp Gly Pro
 385 390 395 400
 Ala Lys Val Tyr Asn Ser Glu Asp Glu Ala Ile Ser Ala Ile Leu Gly
 405 410 415
 Gly Lys Ile Val Glu Gly Asp Val Ile Val Ile Arg Tyr Glu Gly Pro
 420 425 430
 Ser Gly Gly Pro Gly Met Arg Glu Met Leu Ser Pro Thr Ser Ala Ile
 435 440 445
 Cys Gly Met Gly Leu Asp Asp Ser Val Ala Leu Ile Thr Asp Gly Arg
 450 455 460
 Phe Ser Gly Gly Ser Arg Gly Pro Cys Ile Gly His Val Ser Pro Glu
 465 470 475 480
 Ala Ala Ala Gly Gly Val Ile Ala Ala Ile Glu Asn Gly Asp Ile Ile
 485 490 495
 Lys Ile Asp Met Ile Glu Lys Glu Ile Asn Val Asp Leu Asp Glu Ser
 500 505 510
 Val Ile Lys Glu Arg Leu Ser Lys Leu Gly Glu Phe Glu Pro Lys Ile
 515 520 525
 Lys Lys Gly Tyr Leu Ser Arg Tyr Ser Lys Leu Val Ser Ser Ala Asp
 530 535 540
 Glu Gly Ala Val Leu Lys
 545 550

<210> 22
 <211> 1653
 <212> DNA
 <213> Methanococcus mari pal udi s

<400> 22
 atgataagtg ataacgtcaa aaaggaggtt ataagaactc caaaccgagc tcttttaaag 60
 gcttgcggtat atacagacga agacatggaa aaaccattta ttggaattgt aaacagcttt 120
 acagaagttg ttcccggcca cattcactta agaacattat cagaagcggc taaacatggt 180
 gtttatgcaa acggttgaac accatttgaa tttaatacca ttggaatttg cgacgggtatt 240
 gcaatgggcc acgaaggtat gaaatactct ttaccttcaa gagaaattat tgcagacgct 300
 gttgaatcaa tggcaagagc acatggattt gatggtcttg ttttaattcc tacgtgtgat 360
 aaaatcggtc ctggaatgat aatgggtgct ttaagactaa acattccatt tattgtagtt 420
 actggaggac caatgcttc cggagaattc caaggtaaaa aatacgaact tatcagcctt 480
 tttgaagggtg tcggagaata ccaagttgga aaaattactg aagaagagtt aaagtgcatt 540
 gaagactgtg catgttcagg tgctggaagt tgtgcagggc ttactactgc aaacagtatg 600
 gcctgcctta cagaagcttt gggactctct cttccaatgt gtgcaacaac gcatgcagtt 660
 gatgccccaa aagttaggct tgctaaaaaa agtgggtcaa aaattgttga tatggtaaaa 720
 gaagacctaa aaccaacaga catattaaca aaagaagctt ttgaaaatgc tattttagtt 780
 gaccttgac ttggtggatc aacaaacaca acattacaca ttcctgcaat tgcaaatgaa 840
 attgaaaata aattcataac tctcgatgac ttgacaggt taagcgatga agttccacac 900
 attgcatcaa tcaaaccagg tggagaacac tacatgattg atttacacaa tgctggaggt 960
 attcctgcgg tattgaacgt tttaaaagaa aaaattagag atacaaaaac agttgatgga 1020
 agaagcattt tggaaatcgc agaatctgtt aaatacataa attacgacgt tataagaaaa 1080
 gtggaagctc cggttcacga aactgctggt ttaagggttt taaagggaaa tcttgctcca 1140
 aacggttgcg ttgtaaaaat cgggtgcagta catccgaaaa tgtacaaaca cgatggacct 1200
 gcaaaagttt acaattccga agatgaagca atttctgcga tacttggcgg aaaaattgta 1260
 gaaggggacg ttatagtaat cagatacgaa ggaccatcag gaggccctgg aatgagagaa 1320
 atgctctccc caacttcagc aatctgtgga atgggtcttg atgacagcgt tgcattgatt 1380
 actgatggaa gattcagtgg tggaaagtagg ggcccatgta tcggacacgt ttctccagaa 1440
 gctgcagctg gcggagtaat tgctgcaatt gaaaacgggg atatcatcaa aatcgacatg 1500
 attgaaaaag aaataaatgt tgatttagat gaatcagtca ttaaagaaag actctcaaaa 1560
 ctgggagaat ttgagcctaa aatcaaaaaa ggctatttat caagatactc aaaacttgtc 1620
 tcactctgctg acgaaggggc agttttaaaa taa 1653

<210> 23
 <211> 558
 <212> PRT
 <213> Bacillus subtilis

<400> 23

Met Ala Glu Leu Arg Ser Asn Met Ile Thr Gln Gly Ile Asp Arg Ala
 1 5 10 15
 Pro His Arg Ser Leu Leu Arg Ala Ala Gly Val Lys Glu Glu Asp Phe
 20 25 30
 Gly Lys Pro Phe Ile Ala Val Cys Asn Ser Tyr Ile Asp Ile Val Pro
 35 40 45
 Gly His Val His Leu Gln Glu Phe Gly Lys Ile Val Lys Glu Ala Ile
 50 55 60
 Arg Glu Ala Gly Gly Val Pro Phe Glu Phe Asn Thr Ile Gly Val Asp
 65 70 75 80
 Asp Gly Ile Ala Met Gly His Ile Gly Met Arg Tyr Ser Leu Pro Ser
 85 90 95
 Arg Glu Ile Ile Ala Asp Ser Val Glu Thr Val Val Ser Ala His Trp
 100 105 110
 Phe Asp Gly Met Val Cys Ile Pro Asn Cys Asp Lys Ile Thr Pro Gly
 115 120 125
 Met Leu Met Ala Ala Met Arg Ile Asn Ile Pro Thr Ile Phe Val Ser
 130 135 140
 Gly Gly Pro Met Ala Ala Gly Arg Thr Ser Asp Gly Arg Lys Ile Ser
 145 150 155 160
 Leu Ser Ser Val Phe Glu Gly Val Gly Ala Tyr Gln Ala Gly Lys Ile
 165 170 175
 Asn Glu Asn Glu Leu Gln Glu Leu Glu Gln Phe Gly Cys Pro Thr Cys
 180 185 190
 Gly Ser Cys Ser Gly Met Phe Thr Ala Asn Ser Met Asn Cys Leu Ser
 195 200 205
 Glu Ala Leu Gly Leu Ala Leu Pro Gly Asn Gly Thr Ile Leu Ala Thr
 210 215 220
 Ser Pro Glu Arg Lys Glu Phe Val Arg Lys Ser Ala Ala Gln Leu Met
 225 230 235 240
 Glu Thr Ile Arg Lys Asp Ile Lys Pro Arg Asp Ile Val Thr Val Lys
 245 250 255
 Ala Ile Asp Asn Ala Phe Ala Leu Asp Met Ala Leu Gly Gly Ser Thr
 Page 33

260

265

270

Asn Thr Val Leu His Thr Leu Ala Leu Ala Asn Glu Ala Gly Val Glu
 275 280 285

Tyr Ser Leu Glu Arg Ile Asn Glu Val Ala Glu Arg Val Pro His Leu
 290 295 300

Ala Lys Leu Ala Pro Ala Ser Asp Val Phe Ile Glu Asp Leu His Glu
 305 310 315 320

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

Ala Leu His Leu Asp Ala Leu Thr Val Thr Gly Lys Thr Leu Gly Glu
 340 345 350

Thr Ile Ala Gly His Glu Val Lys Asp Tyr Asp Val Ile His Pro Leu
 355 360 365

Asp Gln Pro Phe Thr Glu Lys Gly Gly Leu Ala Val Leu Phe Gly Asn
 370 375 380

Leu Ala Pro Asp Gly Ala Ile Ile Lys Thr Gly Gly Val Gln Asn Gly
 385 390 395 400

Ile Thr Arg His Glu Gly Pro Ala Val Val Phe Asp Ser Gln Asp Glu
 405 410 415

Ala Leu Asp Gly Ile Ile Asn Arg Lys Val Lys Glu Gly Asp Val Val
 420 425 430

Ile Ile Arg Tyr Glu Gly Pro Lys Gly Gly Pro Gly Met Pro Glu Met
 435 440 445

Leu Ala Pro Thr Ser Gln Ile Val Gly Met Gly Leu Gly Pro Lys Val
 450 455 460

Ala Leu Ile Thr Asp Gly Arg Phe Ser Gly Ala Ser Arg Gly Leu Ser
 465 470 475 480

Ile Gly His Val Ser Pro Glu Ala Ala Glu Gly Gly Pro Leu Ala Phe
 485 490 495

Val Glu Asn Gly Asp His Ile Ile Val Asp Ile Glu Lys Arg Ile Leu
 500 505 510

Asp Val Gln Val Pro Glu Glu Glu Trp Glu Lys Arg Lys Ala Asn Trp
 515 520 525

Lys Gly Phe Glu Pro Lys Val Lys Thr Gly Tyr Leu Ala Arg Tyr Ser
 Page 34

530

535

540

Lys Leu Val Thr Ser Ala Asn Thr Gly Gly Ile Met Lys Ile
 545 550 555

<210> 24
 <211> 1677
 <212> DNA
 <213> *Bacillus subtilis*

<400> 24
 atggcagaat tacgcagtaa tatgatcaca caaggaatcg atagagctcc gcaccgcagt 60
 ttgcttcgtg cagcaggggt aaaagaagag gatttcggca agccgtttat tgcggtgtgt 120
 aattcatata ttgatatcgt tcccgggtcat gttcacttgc aggagtttgg gaaaatcgta 180
 aaagaagcaa tcagagaagc agggggcggt ccgtttgaat ttaataccat tggggtagat 240
 gatggcatcg caatggggca tatcggtatg agatattcgc tgccaagccg tgaaattatc 300
 gcagactctg tggaaacggt tgtatccgca cactggtttg acggaatggt ctgtattccg 360
 aactgcgaca aaatcacacc gggaatgctt atggcggcaa tgcgcatcaa cattccgacg 420
 atttttgtca gcggcggacc gatggcggca ggaagaacaa gttacgggcg aaaaatctcc 480
 ctttcctcag tattcgaagg ggtaggcgcc taccaagcag ggaaaatcaa cgaaaacgag 540
 cttcaagaac tagagcagtt cggatgcccc acgtgcgggt cttgctcagg catgtttacg 600
 gcgaactcaa tgaactgtct gtcagaagca cttggtcttg ctttgccggg taatggaacc 660
 attctggcaa catctccgga acgcaaagag tttgtgagaa aatcggctgc gcaattaatg 720
 gaaacgattc gcaaagatat caaacgcgt gatattgtta cagtaaaagc gattgataac 780
 gcgtttgac tcgatatggc gctcggaggt tctacaaata ccgttcttca tacccttgcc 840
 cttgcaaacg aagccggcgt tgaatactct ttagaacgca ttaacgaagt cgctgagcgc 900
 gtgccgcact tggctaagct ggcgcctgca tcggatgtgt ttattgaaga tcttcacgaa 960
 gcgggaggcg tttcagcggc tctgaatgag ctttcgaaga aagaaggagc gcttcattta 1020
 gatgcgctga ctgttacagg aaaaactctt ggagaaacca ttgccggaca tgaagtaaag 1080
 gattatgacg tcattcacc cgtggatcaa ccattcactg aaaagggagg ccttgctgtt 1140
 ttattcggtat atctagctcc ggacggcgct atcattaaaa caggcggcgt acagaatggg 1200
 attacaagac acgaagggcc ggctgtcgta ttcgattctc aggacgaggc gcttgacggc 1260
 attatcaacc gaaaagtaaa agaaggcgac gttgtcatca tcagatacga agggccaaaa 1320
 ggcggacctg gcatgccgga aatgctggcg ccaacatccc aaatcgttgg aatgggactc 1380
 gggccaaaag tggcattgat tacggacgga cgtttttccg gagcctcccg tggcctctca 1440
 atcggccacg tatcacctga ggccgctgag ggccggccgc ttgcctttgt tgaaaacgga 1500
 gaccatatta tcgttgatat tgaaaaacgc atcttgatg tacaagtgcc agaagaagag 1560
 tgggaaaaac gaaaagcgaa ctggaaaggt tttgaaccga aagtgaaaac cggctacctg 1620
 gcacgttatt ctaaacttgt gacaagtgcc aacaccggcg gtattatgaa aatctag 1677

<210> 25
 <211> 547
 <212> PRT
 <213> Lactococcus lactis

<400> 25

```

Met Tyr Thr Val Gly Asp Tyr Leu Leu Asp Arg Leu His Glu Leu Gly
1      5      10      15

Ile Glu Glu Ile Phe Gly Val Pro Gly Asp Tyr Asn Leu Gln Phe Leu
      20      25      30

Asp Gln Ile Ile Ser Arg Glu Asp Met Lys Trp Ile Gly Asn Ala Asn
      35      40      45

Glu Leu Asn Ala Ser Tyr Met Ala Asp Gly Tyr Ala Arg Thr Lys Lys
      50      55      60

Ala Ala Ala Phe Leu Thr Thr Phe Gly Val Gly Glu Leu Ser Ala Ile
65      70      75      80

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

Val Gly Ser Pro Thr Ser Lys Val Gln Asn Asp Gly Lys Phe Val His
      100      105      110

His Thr Leu Ala Asp Gly Asp Phe Lys His Phe Met Lys Met His Glu
      115      120      125

Pro Val Thr Ala Ala Arg Thr Leu Leu Thr Ala Glu Asn Ala Thr Tyr
      130      135      140

Glu Ile Asp Arg Val Leu Ser Gln Leu Leu Lys Glu Arg Lys Pro Val
145      150      155      160

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

Ala Leu Ser Leu Glu Lys Glu Ser Ser Thr Thr Asn Thr Thr Glu Gln
      180      185      190

Val Ile Leu Ser Lys Ile Glu Glu Ser Leu Lys Asn Ala Gln Lys Pro
      195      200      205

Val Val Ile Ala Gly His Glu Val Ile Ser Phe Gly Leu Glu Lys Thr
      210      215      220

Val Thr Gln Phe Val Ser Glu Thr Lys Leu Pro Ile Thr Thr Leu Asn
225      230      235      240

```

Phe Gly Lys Ser Ala Val Asp Glu Ser Leu Pro Ser Phe Leu Gly Ile
 245 250 255
 Tyr Asn Gly Lys Leu Ser Glu Ile Ser Leu Lys Asn Phe Val Glu Ser
 260 265 270
 Ala Asp Phe Ile Leu Met Leu Gly Val Lys Leu Thr Asp Ser Ser Thr
 275 280 285
 Gly Ala Phe Thr His His Leu Asp Glu Asn Lys Met Ile Ser Leu Asn
 290 295 300
 Ile Asp Glu Gly Ile Ile Phe Asn Lys Val Val Glu Asp Phe Asp Phe
 305 310 315 320
 Arg Ala Val Val Ser Ser Leu Ser Glu Leu Lys Gly Ile Glu Tyr Glu
 325 330 335
 Gly Gln Tyr Ile Asp Lys Gln Tyr Glu Glu Phe Ile Pro Ser Ser Ala
 340 345 350
 Pro Leu Ser Gln Asp Arg Leu Trp Gln Ala Val Glu Ser Leu Thr Gln
 355 360 365
 Ser Asn Glu Thr Ile Val Ala Glu Gln Gly Thr Ser Phe Phe Gly Ala
 370 375 380
 Ser Thr Ile Phe Leu Lys Ser Asn Ser Arg Phe Ile Gly Gln Pro Leu
 385 390 395 400
 Trp Gly Ser Ile Gly Tyr Thr Phe Pro Ala Ala Leu Gly Ser Gln Ile
 405 410 415
 Ala Asp Lys Glu Ser Arg His Leu Leu Phe Ile Gly Asp Gly Ser Leu
 420 425 430
 Gln Leu Thr Val Gln Glu Leu Gly Leu Ser Ile Arg Glu Lys Leu Asn
 435 440 445
 Pro Ile Cys Phe Ile Ile Asn Asn Asp Gly Tyr Thr Val Glu Arg Glu
 450 455 460
 Ile His Gly Pro Thr Gln Ser Tyr Asn Asp Ile Pro Met Trp Asn Tyr
 465 470 475 480
 Ser Lys Leu Pro Glu Thr Phe Gly Ala Thr Glu Asp Arg Val Val Ser
 485 490 495
 Lys Ile Val Arg Thr Glu Asn Glu Phe Val Ser Val Met Lys Glu Ala
 500 505 510

Gln Ala Asp Val Asn Arg Met Tyr Trp Ile Glu Leu Val Leu Glu Lys
 515 520 525

Glu Asp Ala Pro Lys Leu Leu Lys Lys Met Gly Lys Leu Phe Ala Glu
 530 535 540

Gln Asn Lys
 545

<210> 26
 <211> 1828
 <212> DNA
 <213> Lactococcus lactis

<400> 26
 tttaaataag tcaatatcgt tgacttattt agaagaaaga gttattcttt aaatgtcaag 60
 ttagttgact aaattaaata taaaatatgg aggaatgtga tgtatacagt aggagattac 120
 ctgttagacc gattacacga gttgggaatt gaagaaattt ttggagttcc tggtgactat 180
 aacttacaat ttttagatca aattatttca cgcgaagata tgaaatggat tggaaatgct 240
 aatgaattaa atgcttctta tatggctgat ggttatgctc gtactaaaaa agctgccgca 300
 tttctcacca catttggagt cggcgaattg agtgcgatca atggactggc aggaagttat 360
 gccgaaaatt taccagtagt agaaattgtt ggttcaccaa cttcaaaagt acaaaatgac 420
 ggaaaatttg tccatcatc actagcagat ggtgatttta aacactttat gaagatgcat 480
 gaacctgtta cagcagcgcg gactttactg acagcagaaa atgccacata tgaaattgac 540
 cgagtacttt ctcaattact aaaagaaaga aaaccagtct atattaactt accagtcgat 600
 gttgctgcag caaaagcaga gaagcctgca ttatctttag aaaaagaaag ctctacaaca 660
 aatacaactg aacaagtgat tttgagtaag attgaagaaa gtttgaaaaa tgcccaaaaa 720
 ccagtagtga ttgcaggaca cgaagtaatt agttttgggt tagaaaaaac ggtaactcag 780
 tttgtttcag aaacaaaact accgattacg acactaaatt ttggtaaaag tgctgttgat 840
 gaatctttgc cctcattttt aggaatatat aacgggaaac tttcagaaat cagtcttaaa 900
 aattttgtgg agtccgcaga ctttatccta atgcttggag tgaagcttac ggactcctca 960
 acaggtgcat tcacacatca tttagatgaa aataaaatga tttactaaa catagatgaa 1020
 ggaataattt tcaataaagt ggtagaagat tttgatttta gagcagtggg ttcttcttta 1080
 tcagaattaa aaggaataga atatgaagga caatatattg ataagcaata tgaagaattt 1140
 attccatcaa gtgctccctt atcacaagac cgtctatggc aggcagttga aagtttgact 1200
 caaagcaatg aaacaatcgt tgctgaacaa ggaacctcat tttttggagc ttcaacaatt 1260
 ttcttaaaat caaatagtcg ttttatttga caacctttat ggggttctat tggatatact 1320
 tttccagcgg ctttaggaag ccaaattgcg gataaagaga gcagacacct tttatttatt 1380
 ggtgatgggt cacttcaact taccgtacaa gaattaggac tatcaatcag agaaaaactc 1440
 aatccaattt gttttatcat aaataatgat ggttatacag ttgaaagaga aatccacgga 1500

cctactcaaa gttataacga cattccaatg tggaattact cgaaattacc agaaacattt 1560
 ggagcaacag aagatcgtgt agtatcaaaa attgttagaa cagagaatga atttgtgtct 1620
 gtcatgaaag aagcccaagc agatgtcaat agaattgtatt ggatagaact agttttggaa 1680
 aaagaagatg cgccaaaatt actgaaaaaa atgggtaaat tatttgctga gcaaaaataaa 1740
 tagatatcaa cggatgatga aaagtaaaat agacaaagtc caataatttt ataaaaagta 1800
 aaaacattag gattttccta atgttttt 1828

<210> 27
 <211> 548
 <212> PRT
 <213> Lactococcus lactis
 <400> 27

Met Tyr Thr Val Gly Asp Tyr Leu Leu Asp Arg Leu His Glu Leu Gly
 1 5 10 15

Ile Glu Glu Ile Phe Gly Val Pro Gly Asp Tyr Asn Leu Gln Phe Leu
 20 25 30

Asp Gln Ile Ile Ser His Lys Asp Met Lys Trp Val Gly Asn Ala Asn
 35 40 45

Glu Leu Asn Ala Ser Tyr Met Ala Asp Gly Tyr Ala Arg Thr Lys Lys
 50 55 60

Ala Ala Ala Phe Leu Thr Thr Phe Gly Val Gly Glu Leu Ser Ala Val
 65 70 75 80

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

Val Gly Ser Pro Thr Ser Lys Val Gln Asn Glu Gly Lys Phe Val His
 100 105 110

His Thr Leu Ala Asp Gly Asp Phe Lys His Phe Met Lys Met His Glu
 115 120 125

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

Glu Ile Asp Arg Val Leu Ser Ala Leu Leu Lys Glu Arg Lys Pro Val
 145 150 155 160

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

Ser Leu Pro Leu Lys Lys Glu Asn Ser Thr Ser Asn Thr Ser Asp Gln
 180 185 190

Glu Ile Leu Asn Lys Ile Gln Glu Ser Leu Lys Asn Ala Lys Lys Pro
 195 200 205
 Ile Val Ile Thr Gly His Glu Ile Ile Ser Phe Gly Leu Glu Lys Thr
 210 215 220
 Val Thr Gln Phe Ile Ser Lys Thr Lys Leu Pro Ile Thr Thr Leu Asn
 225 230 235 240
 Phe Gly Lys Ser Ser Val Asp Glu Ala Leu Pro Ser Phe Leu Gly Ile
 245 250 255
 Tyr Asn Gly Thr Leu Ser Glu Pro Asn Leu Lys Glu Phe Val Glu Ser
 260 265 270
 Ala Asp Phe Ile Leu Met Leu Gly Val Lys Leu Thr Asp Ser Ser Thr
 275 280 285
 Gly Ala Phe Thr His His Leu Asn Glu Asn Lys Met Ile Ser Leu Asn
 290 295 300
 Ile Asp Glu Gly Lys Ile Phe Asn Glu Arg Ile Gln Asn Phe Asp Phe
 305 310 315 320
 Glu Ser Leu Ile Ser Ser Leu Leu Asp Leu Ser Glu Ile Glu Tyr Lys
 325 330 335
 Gly Lys Tyr Ile Asp Lys Lys Gln Glu Asp Phe Val Pro Ser Asn Ala
 340 345 350
 Leu Leu Ser Gln Asp Arg Leu Trp Gln Ala Val Glu Asn Leu Thr Gln
 355 360 365
 Ser Asn Glu Thr Ile Val Ala Glu Gln Gly Thr Ser Phe Phe Gly Ala
 370 375 380
 Ser Ser Ile Phe Leu Lys Ser Lys Ser His Phe Ile Gly Gln Pro Leu
 385 390 395 400
 Trp Gly Ser Ile Gly Tyr Thr Phe Pro Ala Ala Leu Gly Ser Gln Ile
 405 410 415
 Ala Asp Lys Glu Ser Arg His Leu Leu Phe Ile Gly Asp Gly Ser Leu
 420 425 430
 Gln Leu Thr Val Gln Glu Leu Gly Leu Ala Ile Arg Glu Lys Ile Asn
 435 440 445
 Pro Ile Cys Phe Ile Ile Asn Asn Asp Gly Tyr Thr Val Glu Arg Glu
 450 455 460

20130927_CL5646WOPCT_ST25. txt

Ile His Gly Pro Asn Gln Ser Tyr Asn Asp Ile Pro Met Trp Asn Tyr
465 470 475 480

Ser Lys Leu Pro Glu Ser Phe Gly Ala Thr Glu Asp Arg Val Val Ser
485 490 495

Lys Ile Val Arg Thr Glu Asn Glu Phe Val Ser Val Met Lys Glu Ala
500 505 510

Gln Ala Asp Pro Asn Arg Met Tyr Trp Ile Glu Leu Ile Leu Ala Lys
515 520 525

Glu Gly Ala Pro Lys Val Leu Lys Lys Met Gly Lys Leu Phe Ala Glu
530 535 540

Gln Asn Lys Ser
545

<210> 28
<211> 1954
<212> DNA
<213> Lactococcus lactis

<400> 28
ctagagtgtt ctttagtcat aattcactcc ttttattagt ctattatact tgataattca 60
aataagtcaa tatcgttgac ttatttaaag aaaagcgtaa ttctataaat gtcaagttga 120
ttgaccaata tataataaaa tatggaggaa tgcgatgtat acagtaggag attacctatt 180
agaccgatta cacgagttag gaattgaaga aatTTTTGGA gtccctggag actataactt 240
acaattttta gatcaaatta tttccacaa ggatatgaaa tgggtcggaa atgctaataga 300
attaaatgct tcatatatgg ctgatggcta tgctcgtact aaaaaagctg ccgcatttct 360
tacaaccttt ggagtaggtg aattgagtgac agttaatgga ttagcaggaa gttacgccga 420
aaatttacca gtagtagaaa tagtgggatac acctacatca aaagttcaaa atgaaggaaa 480
atttgttcat catagctgg ctgacggtga ttttaaacac tttatgaaaa tgcacgaacc 540
tgttacagca gtcgaactt tactgacagc agaaaatgca accgttgaaa ttgaccgagt 600
actttctgca ctattaaaag aaagaaaacc tgtctatatc aacttaccag ttgatgttgc 660
tgctgcaaaa gcagagaaac cctcactccc tttgaaaaag gaaaactcaa cttcaaatac 720
aagtgaccaa gaaattttga acaaaattca agaaagcttg aaaaatgcca aaaaaccaat 780
cgtgattaca ggacatgaaa taattagttt tggcttagaa aaaacagtca ctcaatttat 840
ttcaaagaca aaactaccta ttacgacatt aaactttggt aaaagttcag ttgatgaagc 900
cctcccttca tttttaggaa tctataatgg tacactctca gagcctaatac ttaaagaatt 960
cgtggaatca gccgacttca tcttgatgct tggagttaaa ctacacagact cttcaacagg 1020
agccttcact catcatttaa atgaaaataa aatgatttca ctgaatatag atgaaggaaa 1080
aatatttaac gaaagaatcc aaaattttga ttttgaatcc ctcatctcct ctctcttaga 1140

```

cctaagcgaa atagaataca aaggaaaata tatcgataaa aagcaagaag actttgttcc 1200
atcaaatgcg cttttatcac aagaccgcct atggcaagca gttgaaaacc taactcaaag 1260
caatgaaaca atcgttgctg aacaagggac atcattcttt ggcgcttcat caattttctt 1320
aaaatcaaag agtcatttta ttggtcaacc cttatgggga tcaattggat atacattccc 1380
agcagcatta ggaagccaaa ttgcagataa agaaagcaga caccttttat ttattggtga 1440
tggttcactt caacttacag tgcaagaatt aggattagca atcagagaaa aaattaatcc 1500
aatttgcttt attatcaata atgatggta tacagtcgaa agagaaattc atggacccaa 1560
tcaaagctac aatgatattc caatgtggaa ttactcaaaa ttaccagaat cgtttgagc 1620
aacagaagat cgagtagtct caaaaatcgt tagaactgaa aatgaatttg tgtctgtcat 1680
gaaagaagct caagcagatc caatagaat gtactggatt gagttaattt tggcaaaaga 1740
agggtcacca aaagtactga aaaaaatggg caaactattt gctgaacaaa ataaatcata 1800
atttataaat agtaaaaaac attaggaaat acctaatggt tttttgttga ctaaataaat 1860
ccctctttat atagaaaacc ttagtttctc aaagacaact taattaagcc tgccaaattg 1920
gaactcgcaa aatgtaatct atcctctgct ccta 1954

```

```

<210> 29
<211> 550
<212> PRT
<213> Salmonella typhimurium
<400> 29

```

```

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

```

```

Gly Cys Gly Ile Gly His Leu Phe Gly Val Pro Gly Asp Tyr Asn Leu
20      25      30

```

```

Gln Phe Leu Asp His Val Ile Asp His Pro Thr Leu Arg Trp Val Gly
35      40      45

```

```

Cys Ala Asn Glu Leu Asn Ala Ala Tyr Ala Ala Asp Gly Tyr Ala Arg
50      55      60

```

```

Met Ser Gly Ala Gly Ala Leu Leu Thr Thr Phe Gly Val Gly Glu Leu
65      70      75      80

```

```

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

```

```

Leu His Ile Val Gly Ala Pro Cys Ser Ala Ala Gln Gln Arg Gly Glu
100     105     110

```

```

Leu Met His His Thr Leu Gly Asp Gly Asp Phe Arg His Phe Tyr Arg
115     120     125

```

20130927_CL5646WOPCT_ST25. txt

Met Ser Gln Ala Ile Ser Ala Ala Ser Ala Ile Leu Asp Glu Gln Asn
130 135 140

Ala Cys Phe Glu Ile Asp Arg Val Leu Gly Glu Met Leu Ala Ala Arg
145 150 155 160

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

Ile Pro Pro Thr Gln Ala Leu Ala Leu Pro Val His Glu Ala Gln Ser
180 185 190

Gly Val Glu Thr Ala Phe Arg Tyr His Ala Arg Gln Cys Leu Met Asn
195 200 205

Ser Arg Arg Ile Ala Leu Leu Ala Asp Phe Leu Ala Gly Arg Phe Gly
210 215 220

Leu Arg Pro Leu Leu Gln Arg Trp Met Ala Glu Thr Pro Ile Ala His
225 230 235 240

Ala Thr Leu Leu Met Gly Lys Gly Leu Phe Asp Glu Gln His Pro Asn
245 250 255

Phe Val Gly Thr Tyr Ser Ala Gly Ala Ser Ser Lys Glu Val Arg Gln
260 265 270

Ala Ile Glu Asp Ala Asp Arg Val Ile Cys Val Gly Thr Arg Phe Val
275 280 285

Asp Thr Leu Thr Ala Gly Phe Thr Gln Gln Leu Pro Ala Glu Arg Thr
290 295 300

Leu Glu Ile Gln Pro Tyr Ala Ser Arg Ile Gly Glu Thr Trp Phe Asn
305 310 315 320

Leu Pro Met Ala Gln Ala Val Ser Thr Leu Arg Glu Leu Cys Leu Glu
325 330 335

Cys Ala Phe Ala Pro Pro Pro Thr Arg Ser Ala Gly Gln Pro Val Arg
340 345 350

Ile Asp Lys Gly Glu Leu Thr Gln Glu Ser Phe Trp Gln Thr Leu Gln
355 360 365

Gln Tyr Leu Lys Pro Gly Asp Ile Ile Leu Val Asp Gln Gly Thr Ala
370 375 380

Ala Phe Gly Ala Ala Ala Leu Ser Leu Pro Asp Gly Ala Glu Val Val
385 390 395 400

20130927_CL5646WOPCT_ST25.txt

Leu Gln Pro Leu Trp Gly Ser Ile Gly Tyr Ser Leu Pro Ala Ala Phe
405 410 415

Gly Ala Gln Thr Ala Cys Pro Asp Arg Arg Val Ile Leu Ile Ile Gly
420 425 430

Asp Gly Ala Ala Gln Leu Thr Ile Gln Glu Met Gly Ser Met Leu Arg
435 440 445

Asp Gly Gln Ala Pro Val Ile Leu Leu Leu Asn Asn Asp Gly Tyr Thr
450 455 460

Val Glu Arg Ala Ile His Gly Ala Ala Gln Arg Tyr Asn Asp Ile Ala
465 470 475 480

Ser Trp Asn Trp Thr Gln Ile Pro Pro Ala Leu Asn Ala Ala Gln Gln
485 490 495

Ala Glu Cys Trp Arg Val Thr Gln Ala Ile Gln Leu Ala Glu Val Leu
500 505 510

Glu Arg Leu Ala Arg Pro Gln Arg Leu Ser Phe Ile Glu Val Met Leu
515 520 525

Pro Lys Ala Asp Leu Pro Glu Leu Leu Arg Thr Val Thr Arg Ala Leu
530 535 540

Glu Ala Arg Asn Gly Gly
545 550

<210> 30
<211> 1653
<212> DNA
<213> Salmonella typhimurium

<400> 30
ttatcccccg ttgcgggctt ccagcgcccg ggtcacggta cgcagtaatt ccggcagatc 60
ggcttttggc aacatcactt caataaatga cagacgttgt gggcgcgcca accgttcgag 120
gacctctgcc agttggatag cctgcgtcac ccgccagcac tccgcctgtt gcgccgcgtt 180
tagcgccggt ggtatctgcg tccagttcca gtcgcgatg tcgttatacc gctgggccgc 240
gccgtgaatg gcgcgctcta cggtatagcc gtcattgttg agcagcagga tgaccggcgc 300
ctgcccgtcg cgtaacatcg agccatctc ctgaatcgtg agctgcgccg cgccatcgcc 360
gataatcaga atcacccgcc gatcgggaca ggcggtttgc gcgcaaacg cggcgggcaa 420
ggaatagccg atagaccccc acagcggctg taacacaact tccgcgccgt caggaagcga 480
cagcgcgga gcgcaaaag ctgctgtccc ctggctgaca aggataatat ctccgggttt 540
gagatactgc tgtaaggttt gccagaagct ttcctgggtc agttctcctt tatcaatccg 600
cactggctgt ccggcggaac gcgtcggcgg cggcgcaaaa gcgcattcca ggcacagttc 660

20130927_CL5646WOPCT_ST25. txt

```

gcgcagcgta gacaccgcct gcgccatcgg gaggttgaac caggtttcgc cgatgcgcga      720
cgcgtaaggc tgaatctcca gcgtgcgttc cgccggtaat tgttgggtaa atccggccgt      780
aagggtatcg acaaaacggg tgccgacgca gataacccta tcggcgctcct ctatggcctg      840
acgcacttct ttgctgctgg cgccagcgct ataggtgcca acgaagttcg ggtgctgttc      900
atcaaaaagc cccttcccca tcagtagtgt cgcatgagcg atgggcgttt ccgccatcca      960
gcgctgcaac agtggtcgta aacaaaaacg cccggcaaga aagtcggcca atagcgcaat     1020
gcgccgactg ttcatacaggc actgacgggc gtgataacga aaggccgtct ccacgccgct     1080
ttgcgcttca tgcacgggca acgccagcgc ctgcgtaggt gggatggccg ttttttcgc     1140
cacatcggcg ggcaacatga tgtatcctgg cctgcgtgcg gcaagcattt caccacacac     1200
gcggtcaatc tcgaaacagg cgttctgttc atctaattt gcgctggcag cggatatcgc     1260
ctgactcatg cgataaaaat gacgaaaatc gccgtcaccg agggatatgt gcatcaattc     1320
gccacgctgc tgcgcagcgc tacagggcgc gccgacgata tgcaagaccg ggacatatcc     1380
cgcgtaactg cccgcgatac cgtaataagc gctaagttct cccacgcca aagggtgag      1440
tagcgctcca gcgcccgaca tgcgcgcata gccgtccgcg gcataagcgg cgttcagctc     1500
attggcgcat cccacccaac gcagggtcgg gtggtcaatc acatggtcaa gaaactgcaa     1560
gttataatcg cccggtacgc caaaaagatg gccaatgccg catcctgcca gtctgtccag     1620
caaatagtcg gccacggtat aggggttttg cat                                  1653

```

<210> 31
 <211> 554
 <212> PRT
 <213> Clostridium acetobutylicum
 <400> 31

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

Glu Leu Gly Ile Arg His Ile Phe Gly Val Pro Gly Asp Tyr Asn Leu
 20 25 30

Ser Phe Leu Asp Tyr Ile Met Glu Tyr Lys Gly Ile Asp Trp Val Gly
 35 40 45

Asn Cys Asn Glu Leu Asn Ala Gly Tyr Ala Ala Asp Gly Tyr Ala Arg
 50 55 60

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

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

Val Lys Ile Thr Gly Ile Pro Thr Ala Lys Val Arg Asp Asn Gly Leu

100

105

110

Tyr Val His His Thr Leu Gly Asp Gly Arg Phe Asp His Phe Phe Glu
 115 120 125

Met Phe Arg Glu Val Thr Val Ala Glu Ala Leu Leu Ser Glu Glu Asn
 130 135 140

Ala Ala Gln Glu Ile Asp Arg Val Leu Ile Ser Cys Trp Arg Gln Lys
 145 150 155 160

Arg Pro Val Leu Ile Asn Leu Pro Ile Asp Val Tyr Asp Lys Pro Ile
 165 170 175

Asn Lys Pro Leu Lys Pro Leu Leu Asp Tyr Thr Ile Ser Ser Asn Lys
 180 185 190

Glu Ala Ala Cys Glu Phe Val Thr Glu Ile Val Pro Ile Ile Asn Arg
 195 200 205

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

Val Gln His Val Leu Lys Asn Leu Ala Glu Lys Thr Gly Phe Pro Val
 225 230 235 240

Ala Thr Leu Ser Met Gly Lys Gly Val Phe Asn Glu Ala His Pro Gln
 245 250 255

Phe Ile Gly Val Tyr Asn Gly Asp Val Ser Ser Pro Tyr Leu Arg Gln
 260 265 270

Arg Val Asp Glu Ala Asp Cys Ile Ile Ser Val Gly Val Lys Leu Thr
 275 280 285

Asp Ser Thr Thr Gly Gly Phe Ser His Gly Phe Ser Lys Arg Asn Val
 290 295 300

Ile His Ile Asp Pro Phe Ser Ile Lys Ala Lys Gly Lys Lys Tyr Ala
 305 310 315 320

Pro Ile Thr Met Lys Asp Ala Leu Thr Glu Leu Thr Ser Lys Ile Glu
 325 330 335

His Arg Asn Phe Glu Asp Leu Asp Ile Lys Pro Tyr Lys Ser Asp Asn
 340 345 350

Gln Lys Tyr Phe Ala Lys Glu Lys Pro Ile Thr Gln Lys Arg Phe Phe
 355 360 365

Glu Arg Ile Ala His Phe Ile Lys Glu Lys Asp Val Leu Leu Ala Glu

370

375

380

Gln Gly Thr Cys Phe Phe Gly Ala Ser Thr Ile Gln Leu Pro Lys Asp
385 390 395 400

Ala Thr Phe Ile Gly Gln Pro Leu Trp Gly Ser Ile Gly Tyr Thr Leu
405 410 415

Pro Ala Leu Leu Gly Ser Gln Leu Ala Asp Gln Lys Arg Arg Asn Ile
420 425 430

Leu Leu Ile Gly Asp Gly Ala Phe Gln Met Thr Ala Gln Glu Ile Ser
435 440 445

Thr Met Leu Arg Leu Gln Ile Lys Pro Ile Ile Phe Leu Ile Asn Asn
450 455 460

Asp Gly Tyr Thr Ile Glu Arg Ala Ile His Gly Arg Glu Gln Val Tyr
465 470 475 480

Asn Asn Ile Gln Met Trp Arg Tyr His Asn Val Pro Lys Val Leu Gly
485 490 495

Pro Lys Glu Cys Ser Leu Thr Phe Lys Val Gln Ser Glu Thr Glu Leu
500 505 510

Glu Lys Ala Leu Leu Val Ala Asp Lys Asp Cys Glu His Leu Ile Phe
515 520 525

Ile Glu Val Val Met Asp Arg Tyr Asp Lys Pro Glu Pro Leu Glu Arg
530 535 540

Leu Ser Lys Arg Phe Ala Asn Gln Asn Asn
545 550

<210> 32
<211> 1665
<212> DNA
<213> Clostridium acetobutylicum

<400> 32
ttgaagagtg aatacacaat tggaagatat ttgttagacc gtttatcaga gttgggtatt 60
cggcatatct ttggtgtacc tggagattac aatctatcct ttttagacta tataatggag 120
taciaaggga tagattgggt tggaattgc aatgaattga atgctgggta tgctgctgat 180
ggatatgcaa gaataaatgg aattggagcc atacttaca cattedgtgt tggagaatta 240
agtgccatta acgcaattgc tggggcatac gctgagcaag ttccagttgt taaaattaca 300
ggtatcccca cagcaaaagt tagggacaat ggattatatg tacaccacac attaggtgac 360
ggaaggtttg atcacttttt tgaaatgttt agagaagtaa cagttgctga ggcattacta 420
agcgaagaaa atgcagcaca agaaattgat cgtgttctta ttcatgctg gagacaaaaa 480

20130927_CL5646WOPCT_ST25. txt

cgtcctgttc ttataaattt accgattgat gtatatgata aaccaattaa caaaccatta	540
aagccattac tcgattatac ttttcaagt aacaaagagg ctgcatgtga atttgttaca	600
gaaatagtac ctataataaa tagggcaaaa aagcctgtta ttcttgcaga ttatggagta	660
tatcgttacc aagttcaaca tgtgcttaaa aacttggccg aaaaaaccgg atttcctgtg	720
gctacactaa gtatgggaaa aggtgttttc aatgaagcac accctcaatt tatttgggtgtt	780
tataatgggtg atgtaagttc tccttattta aggagcgag ttgatgaagc agactgcatt	840
attagcgttg gtgtaaaatt gacggattca accacagggg gatcttctca tggattttct	900
aaaaggaatg taattcacat tgatcctttt tcaataaagg caaaaggtaa aaaatatgca	960
cctattacga tgaaagatgc tttacagaa ttaacaagta aaattgagca tagaaacttt	1020
gaggatttag atataaagcc ttacaaatca gataatcaaa agtattttgc aaaagagaag	1080
ccaattacac aaaaacgttt ttttgagcgt attgctcact ttataaaaga aaaagatgta	1140
ttatttagcag aacagggtac atgctttttt ggtgctcaa ccatacaact acccaaagat	1200
gcaactttta ttggtcaacc tttatgggga tctattggat acacacttcc tgctttatta	1260
ggttcacaat tagctgatca aaaaaggcgt aatattcttt taattgggga tgggtgcattt	1320
caaatgacag cacaagaaat ttcaacaatg cttcgtttac aaatcaaacc tattattttt	1380
ttaattaata acgatggta tacaattgaa cgtgctattc atggtagaga acaagtatat	1440
aacaatattc aaatgtggcg atatcataat gttccaaagg ttttaggtcc taaagaatgc	1500
agcttaacct ttaaagtaca aagtgaact gaactgaaa aggctctttt agtggcagat	1560
aaggatttg aacatttgat ttttatagaa gttgttatgg atcgttatga taaacccgag	1620
cctttagaac gtctttcgaa acgttttgca aatcaaaata attag	1665

<210> 33
 <211> 1641
 <212> DNA
 <213> Clostridium acetobutylicum

<400> 33	
atgaacaac gtatcgggca atacttgatc gatgccctac acgttaatgg tgtcgataag	60
atctttggag tcccagggtga tttcacttta gccttttttg acgatatcat aagacatgac	120
aacgtggaat ggggtgggaaa tactaatgag ttgaacgccg cttacgccgc tgatggttac	180
gctagagtta atggattagc cgctgtatct accacttttg gggttggcga gttatctgct	240
gtgaatggta ttgctggaag ttacgcagag cgtgttcctg taatcaaaat ctcaggcggt	300
ccttcatcag ttgctcaaca agagggtaga tatgtccacc attcattggg tgaaggaatc	360
tttgattcat attcaaagat gtacgctcac ataaccgcaa caactacaat cttatccgtt	420
gacaacgcag tcgacgaaat tgatagagtt attcattgtg ctttgaagga aaagaggcca	480
gtgcatattc atttgcctat tgacgtagcc ttaactgaga ttgaaatccc tcatgcacca	540
aaagtttaca cacacgaatc ccagaacgtc gatgcttaca ttcaagctgt tgagaaaaag	600

20130927_CL5646WOPCT_ST25. txt

ttaatgtctg caaaacaacc agtaatcata gcaggtcatg aaatcaattc attcaagttg	660
cacgaacaac tggacagtt tgtcaatcag acaaacatcc ctgttgacac actttccttg	720
ggtaagtctg ctttcaatga agagaatgaa cattaccttg gtatctacga tggcaaaatc	780
gcaaaggaaa atgtgagaga gtacgtcgac aatgctgatg tcatattgaa cataggtgcc	840
aaactgactg attctgctac agctggattt tcctacaagt tcgatacaaa caacataatc	900
tacattaacc ataatgactt caaagctgaa gatgtgattt ctgataatgt ttcactgatt	960
gatcttgtga atggcctgaa ttctattgac tatagaaatg aaacacacta cccatcttat	1020
caaagatctg atatgaaata cgaattgaat gacgcaccac ttacacaatc taactatttc	1080
aaaatgatga acgcttttct agaaaaagat gacatcctac tagctgaaca aggtacatcc	1140
tttttcggcg catatgactt atccctatac aagggaaatc agtttatcgg tcagccttta	1200
tgggggtcaa tagggatatac ttttccatct ttactaggaa gtcaactagc agacatgcat	1260
aggagaaaca ttttgcttat aggcgatggt agtttacaac ttactgttca agccctaagt	1320
acaatgatta gaaaggatat caaaccaatc attttcgtta tcaataacga cggttacacc	1380
gtcgaaagac ttatccacgg catggaagag ccatacaatg atatccaaat gtggaactac	1440
aagcaattgc cagaagtatt tgggtgaaaa gatactgtaa aagttcatga tgctaaaacc	1500
tccaacgaac tgaaaactgt aatggattct gttaaagcag acaaagatca catgcatttc	1560
attgaagtgc atatggcagt agaggacgcc ccaaagaagt tgattgatat agctaaagcc	1620
tttagtgatg ctaacaagta a	1641

<210> 34
 <211> 1647
 <212> DNA
 <213> *Listeria grayi*

<400> 34 atgtacaccg tcggccaata cttagtagac cgcttagaag agatcggcat cgataagggtt	60
tttgggtgtcc cgggtgacta caacctgacc tttttggact acatccagaa ccacgaagggt	120
ctgagctggc aaggtaatac gaatgaactg aatgccgcgt acgcagctga tggctatgct	180
cgtgaacgcg gtgttagcgc tttggtcacg accttcggcg ttggtagact gtccgcaatc	240
aatggcaccg caggtagctt cgcggagcaa gttccggtga ttcatatcgt gggcagccc	300
accatgaatg ttacagagcaa caagaaactg gttcatcaca gcctgggtat gggcaacttt	360
cacaacttca gcgagatggc gaaagaagtc accgccgcaa ccacgatgct gacggaagag	420
aatgcggcgt cggagattga tcgtgttctg gaaaccgccc tgctggagaa acgcccagtg	480
tacatcaatc tgccgatcga cattgctcac aaggcgatcg tcaagccggc gaaagccctg	540
caaaccgaga agagctctgg cgagcgtgag gcacaactgg cggagatcat tctgagccat	600
ctggagaagg ctgcacagcc gattgtgatt gcgggtcacg agatcgcgcg cttccagatc	660
cgtgagcgtt tcgagaattg gattaatcaa acgaaactgc cggtgaccaa tctggcctac	720
ggcaagggtg gcttcaacga agaaaacgag catttcattg gtacctatta tcctgcattt	780

```

agcgataaga acgtgctgga ctacgtggat aactccgact ttgtcctgca ctttggtggt      840
aaaatcattg ataacagcac ctccagcttc tccaagggt tcaaaaccga gaacaccctg      900
actgcggcga acgatatcat tatgctgccg gacggtagca cgtattctgg tattagcctg      960
aatggcctgc tggccgagct ggaaaaactg aatttcacgt ttgccgacac cgcagcaaag     1020
caggcggagt tggcgggtgt tgagccgcag gctgaaaccc cgttgaaaca ggaccgtttt     1080
caccaggcgg tgatgaatth tctgcaagct gacgatgtcc tggttacgga acagggcacc     1140
tcttcttttg gcttgatgct ggcgctctg aaaaagggtg tgaacttgat ctcgcaaacg     1200
ctgtggggta gcattggtta cacgttgccg gcgatgattg gtagccaaat tgcggcaccg     1260
gagcgtcgtc atatcctgag cattggtgat ggtagctttc agctgactgc gcaggaaatg     1320
agcaccattt tccgtgagaa actgaccca gtcatttca tcattaacaa tgatggctat     1380
accgttgagc gtgcgatcca tggcgaagat gaaagctata acgacattcc gacgtggaac     1440
ttgcaactgg tggcggaaac cttcgggtgg gacgccgaaa ccgtcgacac tcacaatgtg     1500
ttcacggaga ctgatttcgc caacaccctg gcggcaattg acgcgacgcc gcagaaagca     1560
cacgttggtg aagttcacat ggaacaaatg gatatgccgg agagcctgcg ccagatcggt     1620
ctggcactgt ccaagcagaa tagctaa                                           1647

```

```

<210> 35
<211> 312
<212> PRT
<213> Saccharomyces cerevisiae

```

```
<400> 35
```

```
Met Pro Ala Thr Leu Lys Asn Ser Ser Ala Thr Leu Lys Leu Asn Thr
1          5          10          15
```

```
Gly Ala Ser Ile Pro Val Leu Gly Phe Gly Thr Trp Arg Ser Val Asp
          20          25          30
```

```
Asn Asn Gly Tyr His Ser Val Ile Ala Ala Leu Lys Ala Gly Tyr Arg
          35          40          45
```

```
His Ile Asp Ala Ala Ala Ile Tyr Leu Asn Glu Glu Glu Val Gly Arg
          50          55          60
```

```
Ala Ile Lys Asp Ser Gly Val Pro Arg Glu Glu Ile Phe Ile Thr Thr
65          70          75          80
```

```
Lys Leu Trp Gly Thr Glu Gln Arg Asp Pro Glu Ala Ala Leu Asn Lys
          85          90          95
```

```
Ser Leu Lys Arg Leu Gly Leu Asp Tyr Val Asp Leu Tyr Leu Met His
          100          105          110
```

```
Trp Pro Val Pro Leu Lys Thr Asp Arg Val Thr Asp Gly Asn Val Leu
Page 50
```

115

120

125

Cys Ile Pro Thr Leu Glu Asp Gly Thr Val Asp Ile Asp Thr Lys Glu
130 135 140

Trp Asn Phe Ile Lys Thr Trp Glu Leu Met Gln Glu Leu Pro Lys Thr
145 150 155 160

Gly Lys Thr Lys Ala Val Gly Val Ser Asn Phe Ser Ile Asn Asn Ile
165 170 175

Lys Glu Leu Leu Glu Ser Pro Asn Asn Lys Val Val Pro Ala Thr Asn
180 185 190

Gln Ile Glu Ile His Pro Leu Leu Pro Gln Asp Glu Leu Ile Ala Phe
195 200 205

Cys Lys Glu Lys Gly Ile Val Val Glu Ala Tyr Ser Pro Phe Gly Ser
210 215 220

Ala Asn Ala Pro Leu Leu Lys Glu Gln Ala Ile Ile Asp Met Ala Lys
225 230 235 240

Lys His Gly Val Glu Pro Ala Gln Leu Ile Ile Ser Trp Ser Ile Gln
245 250 255

Arg Gly Tyr Val Val Leu Ala Lys Ser Val Asn Pro Glu Arg Ile Val
260 265 270

Ser Asn Phe Lys Ile Phe Thr Leu Pro Glu Asp Asp Phe Lys Thr Ile
275 280 285

Ser Asn Leu Ser Lys Val His Gly Thr Lys Arg Val Val Asp Met Lys
290 295 300

Trp Gly Ser Phe Pro Ile Phe Gln
305 310

<210> 36
<211> 939
<212> DNA
<213> Saccharomyces cerevisiae

<400> 36
atgcctgcta cgtaaagaa ttcttctgct acattaaaac taaatactgg tgcctccatt 60
ccagtgttgg gtttcggcac ttggcgttcc gttgacaata acggttacca ttctgtaatt 120
gcagctttga aagctggata cagacacatt gatgctgcgg ctatctatit gaatgaagaa 180
gaagtggca gggctattaa agattccgga gtccctcgtg aggaaatit tattactact 240
aagctttggg gtacggaaca acgtgatccg gaagctgctc taaacaagtc ttgaaaaga 300
ctaggcttgg attatgttga cctatatctg atgcattggc cagtgcctit gaaaaccgac 360

```

agagttactg atggtaacgt tctgtgcatt ccaacattag aagatggcac tgttgacatc      420
gatactaagg aatggaattt tatcaagacg tgggagttga tgcaagagtt gccaaagacg      480
ggcaaaaacta aagccgttgg tgtctctaatt ttttctatta acaacattaa agaattatta      540
gaatctccaa ataacaaggt ggtaccagct actaatcaaa ttgaaattca tccattgcta      600
ccacaagacg aattgattgc cttttgtaag gaaaagggtta ttgttggtga agcctactca      660
ccatttggga gtgctaattgc tcctttacta aaagagcaag caattattga tatggctaaa      720
aagcacggcg ttgagccagc acagcttatt atcagttgga gtattcaaag aggctacggt      780
gttctggcca aatcggttaa tcctgaaaga attgtatcca attttaagat tttcactctg      840
cctgaggatg atttcaagac tattagtaac ctatccaaag tgcattggtac aaagagagtc      900
gttgatatga agtggggatc cttcccaatt ttccaatga                               939

```

```

<210> 37
<211> 360
<212> PRT
<213> Saccharomyces cerevisiae
<400> 37

```

```

Met Ser Tyr Pro Glu Lys Phe Glu Gly Ile Ala Ile Gln Ser His Glu
1          5          10          15

Asp Trp Lys Asn Pro Lys Lys Thr Lys Tyr Asp Pro Lys Pro Phe Tyr
20          25          30

Asp His Asp Ile Asp Ile Lys Ile Glu Ala Cys Gly Val Cys Gly Ser
35          40          45

Asp Ile His Cys Ala Ala Gly His Trp Gly Asn Met Lys Met Pro Leu
50          55          60

Val Val Gly His Glu Ile Val Gly Lys Val Val Lys Leu Gly Pro Lys
65          70          75          80

Ser Asn Ser Gly Leu Lys Val Gly Gln Arg Val Gly Val Gly Ala Gln
85          90          95

Val Phe Ser Cys Leu Glu Cys Asp Arg Cys Lys Asn Asp Asn Glu Pro
100         105         110

Tyr Cys Thr Lys Phe Val Thr Thr Tyr Ser Gln Pro Tyr Glu Asp Gly
115         120         125

Tyr Val Ser Gln Gly Gly Tyr Ala Asn Tyr Val Arg Val His Glu His
130         135         140

Phe Val Val Pro Ile Pro Glu Asn Ile Pro Ser His Leu Ala Ala Pro
145         150         155         160

```

20130927_CL5646WOPCT_ST25. txt

Leu Leu Cys Gly Gly Leu Thr Val Tyr Ser Pro Leu Val Arg Asn Gly
165 170 175

Cys Gly Pro Gly Lys Lys Val Gly Ile Val Gly Leu Gly Gly Ile Gly
180 185 190

Ser Met Gly Thr Leu Ile Ser Lys Ala Met Gly Ala Glu Thr Tyr Val
195 200 205

Ile Ser Arg Ser Ser Arg Lys Arg Glu Asp Ala Met Lys Met Gly Ala
210 215 220

Asp His Tyr Ile Ala Thr Leu Glu Glu Gly Asp Trp Gly Glu Lys Tyr
225 230 235 240

Phe Asp Thr Phe Asp Leu Ile Val Val Cys Ala Ser Ser Leu Thr Asp
245 250 255

Ile Asp Phe Asn Ile Met Pro Lys Ala Met Lys Val Gly Gly Arg Ile
260 265 270

Val Ser Ile Ser Ile Pro Glu Gln His Glu Met Leu Ser Leu Lys Pro
275 280 285

Tyr Gly Leu Lys Ala Val Ser Ile Ser Tyr Ser Ala Leu Gly Ser Ile
290 295 300

Lys Glu Leu Asn Gln Leu Leu Lys Leu Val Ser Glu Lys Asp Ile Lys
305 310 315 320

Ile Trp Val Glu Thr Leu Pro Val Gly Glu Ala Gly Val His Glu Ala
325 330 335

Phe Glu Arg Met Glu Lys Gly Asp Val Arg Tyr Arg Phe Thr Leu Val
340 345 350

Gly Tyr Asp Lys Glu Phe Ser Asp
355 360

<210> 38
<211> 1083
<212> DNA
<213> Saccharomyces cerevisiae

<400> 38
ctagtctgaa aattctttgt cgtagccgac taaggtaa atctatatctaa cgtcaccctt 60
ttccatcctt tcgaaggctt catggacgcc ggcttcacca acaggtaatg tttccaccca 120
aatittgata tctttttcag agactaat ttaagagttgg ttcaattctt tgatggaacc 180
taaagcactg taagaaatgg agacagcctt taagccatat ggcttttagcg ataacatttc 240
gtgttgttct ggtatagaga ttgagacaat tctaccacca accttcatag cctttggcat 300

```

aatgttgaag tcaatgtcgg taagggagga agcacagact acaatcaggt cgaaggtgtc      360
aaagtacttt tcacccaat caccttcttc taatgtagca atgtagtgat cggcgcccat      420
cttcattgca tcttctcttt ttctcgaaga acgagaaata acatacgtct ctgcccccat      480
ggctttggaa atcaatgtac ccatactgcc gataccacca agaccaacta taccaacttt      540
tttacctgga ccgcaaccgt tacgaaccaa tggagagtac acagtcaaac caccacataa      600
tagtggagca gccaaatgtg atggaatatt ctctgggata ggcaccacaa aatgttcatg      660
aactctgacg tagtttgcag agccaccctg cgacacatag ccgtcttcat aaggctgact      720
gtatgtggta acaaacttgg tgcagtatgg ttcatatca ttcttacaac ggtcacattc      780
caagcatgaa aagacttgag cacctacacc aacacgttga ccgactttca acccactgtt      840
tgacttgggc cctagcttga caactttacc aacgatttca tgaccaacga ctagcggcat      900
cttcatattg cccaatgac cagctgcaca atgaatatca ctaccgcaga caccacatgc      960
ttcgatctta atgtcaatgt catgatcgta aaatggtttt gggtcatact ttgtcttctt     1020
tgggtttttc caatcttcgt gtgattgaat agcgatacct tcaaatttct caggataaga     1080
cat                                                                    1083

```

```

<210> 39
<211> 387
<212> PRT
<213> Escherichia coli
<400> 39

```

```

Met Asn Asn Phe Asn Leu His Thr Pro Thr Arg Ile Leu Phe Gly Lys
1          5          10
Gly Ala Ile Ala Gly Leu Arg Glu Gln Ile Pro His Asp Ala Arg Val
20          25          30
Leu Ile Thr Tyr Gly Gly Gly Ser Val Lys Lys Thr Gly Val Leu Asp
35          40          45
Gln Val Leu Asp Ala Leu Lys Gly Met Asp Val Leu Glu Phe Gly Gly
50          55          60
Ile Glu Pro Asn Pro Ala Tyr Glu Thr Leu Met Asn Ala Val Lys Leu
65          70          75          80
Val Arg Glu Gln Lys Val Thr Phe Leu Leu Ala Val Gly Gly Gly Ser
85          90          95
Val Leu Asp Gly Thr Lys Phe Ile Ala Ala Ala Ala Asn Tyr Pro Glu
100         105         110
Asn Ile Asp Pro Trp His Ile Leu Gln Thr Gly Gly Lys Glu Ile Lys
115         120         125

```


20130927_CL5646W0PCT_ST25. txt

Ser Ala Ile Pro Met Gly Cys Val Leu Thr Leu Pro Ala Thr Gly Ser
130 135 140

Glu Ser Asn Ala Gly Ala Val Ile Ser Arg Lys Thr Thr Gly Asp Lys
145 150 155 160

Gln Ala Phe His Ser Ala His Val Gln Pro Val Phe Ala Val Leu Asp
165 170 175

Pro Val Tyr Thr Tyr Thr Leu Pro Pro Arg Gln Val Ala Asn Gly Val
180 185 190

Val Asp Ala Phe Val His Thr Val Glu Gln Tyr Val Thr Lys Pro Val
195 200 205

Asp Ala Lys Ile Gln Asp Arg Phe Ala Glu Gly Ile Leu Leu Thr Leu
210 215 220

Ile Glu Asp Gly Pro Lys Ala Leu Lys Glu Pro Glu Asn Tyr Asp Val
225 230 235 240

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

Gly Ala Gly Val Pro Gln Asp Trp Ala Thr His Met Leu Gly His Glu
260 265 270

Leu Thr Ala Met His Gly Leu Asp His Ala Gln Thr Leu Ala Ile Val
275 280 285

Leu Pro Ala Leu Trp Asn Glu Lys Arg Asp Thr Lys Arg Ala Lys Leu
290 295 300

Leu Gln Tyr Ala Glu Arg Val Trp Asn Ile Thr Glu Gly Ser Asp Asp
305 310 315 320

Glu Arg Ile Asp Ala Ala Ile Ala Ala Thr Arg Asn Phe Phe Glu Gln
325 330 335

Leu Gly Val Pro Thr His Leu Ser Asp Tyr Gly Leu Asp Gly Ser Ser
340 345 350

Ile Pro Ala Leu Leu Lys Lys Leu Glu Glu His Gly Met Thr Gln Leu
355 360 365

Gly Glu Asn His Asp Ile Thr Leu Asp Val Ser Arg Arg Ile Tyr Glu
370 375 380

Ala Ala Arg
385

<210> 40
 <211> 387
 <212> PRT
 <213> Escherichia coli
 <400> 40
 Met Asn Asn Phe Asn Leu His Thr Pro Thr Arg Ile Leu Phe Gly Lys
 1 5 10 15
 Gly Ala Ile Ala Gly Leu Arg Glu Gln Ile Pro His Asp Ala Arg Val
 20 25 30
 Leu Ile Thr Tyr Gly Gly Gly Ser Val Lys Lys Thr Gly Val Leu Asp
 35 40 45
 Gln Val Leu Asp Ala Leu Lys Gly Met Asp Val Leu Glu Phe Gly Gly
 50 55 60
 Ile Glu Pro Asn Pro Ala Tyr Glu Thr Leu Met Asn Ala Val Lys Leu
 65 70 75 80
 Val Arg Glu Gln Lys Val Thr Phe Leu Leu Ala Val Gly Gly Gly Ser
 85 90 95
 Val Leu Asp Gly Thr Lys Phe Ile Ala Ala Ala Ala Asn Tyr Pro Glu
 100 105 110
 Asn Ile Asp Pro Trp His Ile Leu Gln Thr Gly Gly Lys Glu Ile Lys
 115 120 125
 Ser Ala Ile Pro Met Gly Cys Val Leu Thr Leu Pro Ala Thr Gly Ser
 130 135 140
 Glu Ser Asn Ala Gly Ala Val Ile Ser Arg Lys Thr Thr Gly Asp Lys
 145 150 155 160
 Gln Ala Phe His Ser Ala His Val Gln Pro Val Phe Ala Val Leu Asp
 165 170 175
 Pro Val Tyr Thr Tyr Thr Leu Pro Pro Arg Gln Val Ala Asn Gly Val
 180 185 190
 Val Asp Ala Phe Val His Thr Val Glu Gln Tyr Val Thr Lys Pro Val
 195 200 205
 Asp Ala Lys Ile Gln Asp Arg Phe Ala Glu Gly Ile Leu Leu Thr Leu
 210 215 220
 Ile Glu Asp Gly Pro Lys Ala Leu Lys Glu Pro Glu Asn Tyr Asp Val
 225 230 235 240

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

Gly Ala Gly Val Pro Gln Asp Trp Ala Thr His Met Leu Gly His Glu
260 265 270

Leu Thr Ala Met His Gly Leu Asp His Ala Gln Thr Leu Ala Ile Val
275 280 285

Leu Pro Ala Leu Trp Asn Glu Lys Arg Asp Thr Lys Arg Ala Lys Leu
290 295 300

Leu Gln Tyr Ala Glu Arg Val Trp Asn Ile Thr Glu Gly Ser Asp Asp
305 310 315 320

Glu Arg Ile Asp Ala Ala Ile Ala Ala Thr Arg Asn Phe Phe Glu Gln
325 330 335

Leu Gly Val Pro Thr His Leu Ser Asp Tyr Gly Leu Asp Gly Ser Ser
340 345 350

Ile Pro Ala Leu Leu Lys Lys Leu Glu Glu His Gly Met Thr Gln Leu
355 360 365

Gly Glu Asn His Asp Ile Thr Leu Asp Val Ser Arg Arg Ile Tyr Glu
370 375 380

Ala Ala Arg
385

<210> 41
<211> 389
<212> PRT
<213> Clostridium acetobutylicum

<400> 41

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

Gly Lys Ile Asp Val Ile Gly Glu Glu Ile Lys Lys Tyr Gly Ser Arg
20 25 30

Val Leu Ile Val Tyr Gly Gly Gly Ser Ile Lys Arg Asn Gly Ile Tyr
35 40 45

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

Leu Ser Gly Val Glu Pro Asn Pro Arg Ile Thr Thr Val Lys Lys Gly
65 70 75 80

Ile Glu Ile Cys Arg Glu Asn Asn Val Asp Leu Val Leu Ala Ile Gly
Page 57

Gly Gly Ser Ala Ile Asp Cys Ser Lys Val Ile Ala Ala Gly Val Tyr
 100 105 110
 Tyr Asp Gly Asp Thr Trp Asp Met Val Lys Asp Pro Ser Lys Ile Thr
 115 120 125
 Lys Val Leu Pro Ile Ala Ser Ile Leu Thr Leu Ser Ala Thr Gly Ser
 130 135 140
 Glu Met Asp Gln Ile Ala Val Ile Ser Asn Met Glu Thr Asn Glu Lys
 145 150 155 160
 Leu Gly Val Gly His Asp Asp Met Arg Pro Lys Phe Ser Val Leu Asp
 165 170 175
 Pro Thr Tyr Thr Phe Thr Val Pro Lys Asn Gln Thr Ala Ala Gly Thr
 180 185 190
 Ala Asp Ile Met Ser His Thr Phe Glu Ser Tyr Phe Ser Gly Val Glu
 195 200 205
 Gly Ala Tyr Val Gln Asp Gly Ile Ala Glu Ala Ile Leu Arg Thr Cys
 210 215 220
 Ile Lys Tyr Gly Lys Ile Ala Met Glu Lys Thr Asp Asp Tyr Glu Ala
 225 230 235 240
 Arg Ala Asn Leu Met Trp Ala Ser Ser Leu Ala Ile Asn Gly Leu Leu
 245 250 255
 Ser Leu Gly Lys Asp Arg Lys Trp Ser Cys His Pro Met Glu His Glu
 260 265 270
 Leu Ser Ala Tyr Tyr Asp Ile Thr His Gly Val Gly Leu Ala Ile Leu
 275 280 285
 Thr Pro Asn Trp Met Glu Tyr Ile Leu Asn Asp Asp Thr Leu His Lys
 290 295 300
 Phe Val Ser Tyr Gly Ile Asn Val Trp Gly Ile Asp Lys Asn Lys Asp
 305 310 315 320
 Asn Tyr Glu Ile Ala Arg Glu Ala Ile Lys Asn Thr Arg Glu Tyr Phe
 325 330 335
 Asn Ser Leu Gly Ile Pro Ser Lys Leu Arg Glu Val Gly Ile Gly Lys
 340 345 350
 Asp Lys Leu Glu Leu Met Ala Lys Gln Ala Val Arg Asn Ser Gly Gly

355

360

365

Thr Ile Gly Ser Leu Arg Pro Ile Asn Ala Glu Asp Val Leu Glu Ile
 370 375 380

Phe Lys Lys Ser Tyr
 385

<210> 42
 <211> 1170
 <212> DNA
 <213> Clostridium acetobutylicum

<400> 42
 ttaataagat tttttaaata tctcaagaac atcctctgca tttattggtc ttaaacttcc 60
 tattgttcct ccagaatttc taacagcttg ctttgccatt agttctagtt tatcttttcc 120
 tattccaact tctctaagct ttgaaggaat acccaatgaa ttaaagtatt ctctcgtatt 180
 tttaatagcc tctcgtgcta tttcatagtt atctttgttc ttgtctattc cccaaacatt 240
 tattccataa gaaacaaatt tatgaagtgt atcgtcattt agaatatatt ccatccaatt 300
 aggtgttaaa attgcaagtc ctacaccatg tgttatatca taatatgcac ttaactcgtg 360
 ttccatagga tgacaactcc attttctatc cttaccaagt gataatagac cttttatagc 420
 taaacttgaa gccacatca aattagctct agcctcgtaa tcatcagtct tctccattgc 480
 tatttttcca tactttatac atgttcttaa gattgcttct gctataccgt cctgcacata 540
 agcaccttca acaccactaa agtaagattc aaaggtgtga ctcataatgt cagctgtttcc 600
 cgctgctgtt tgatttttag gtactgtaaa agtatatgta ggatctaaca ctgaaaattt 660
 aggtctcata tcatcatgtc ctactccaag cttttcatta gtctccatat ttgaaattac 720
 tgcaatttga tccatttcag accctgttgc tgaaagagta agtatacttg caattggaag 780
 aactttagtt attttagatg gatctttaac catgtcccat gtatcgccat cataataaac 840
 tccagctgca attaccttag aacagtctat tgcacttcct ccccctattg ctaataactaa 900
 atccacatta ttttctctac atatttctat gccttttttt actgttggtta tcctaggatt 960
 tggctctact cctgaaagtt catagaaagc tatattgttt tcttttaata tagctgttgc 1020
 tctatcatat ataccgttcc tttttatact tcctccgcca taaactataa gcactcttga 1080
 gccatatttc ttaatttctt ctccaattac gtctattttt ctttttccaa aaaaaacttt 1140
 agttggtatt gaataatcaa aacttagcat 1170

<210> 43
 <211> 390
 <212> PRT
 <213> Clostridium acetobutylicum

<400> 43

Met Val Asp Phe Glu Tyr Ser Ile Pro Thr Arg Ile Phe Phe Gly Lys
 1 5 10 15

20130927_CL5646W0PCT_ST25. txt

Asp Lys Ile Asn Val Leu Gly Arg Glu Leu Lys Lys Tyr Gly Ser Lys
20 25 30

Val Leu Ile Val Tyr Gly Gly Gly Ser Ile Lys Arg Asn Gly Ile Tyr
35 40 45

Asp Lys Ala Val Ser Ile Leu Glu Lys Asn Ser Ile Lys Phe Tyr Glu
50 55 60

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

Val Lys Ile Cys Arg Glu Asn Gly Val Glu Val Val Leu Ala Ile Gly
85 90 95

Gly Gly Ser Ala Ile Asp Cys Ala Lys Val Ile Ala Ala Ala Cys Glu
100 105 110

Tyr Asp Gly Asn Pro Trp Asp Ile Val Leu Asp Gly Ser Lys Ile Lys
115 120 125

Arg Val Leu Pro Ile Ala Ser Ile Leu Thr Ile Ala Ala Thr Gly Ser
130 135 140

Glu Met Asp Thr Trp Ala Val Ile Asn Asn Met Asp Thr Asn Glu Lys
145 150 155 160

Leu Ile Ala Ala His Pro Asp Met Ala Pro Lys Phe Ser Ile Leu Asp
165 170 175

Pro Thr Tyr Thr Tyr Thr Val Pro Thr Asn Gln Thr Ala Ala Gly Thr
180 185 190

Ala Asp Ile Met Ser His Ile Phe Glu Val Tyr Phe Ser Asn Thr Lys
195 200 205

Thr Ala Tyr Leu Gln Asp Arg Met Ala Glu Ala Leu Leu Arg Thr Cys
210 215 220

Ile Lys Tyr Gly Gly Ile Ala Leu Glu Lys Pro Asp Asp Tyr Glu Ala
225 230 235 240

Arg Ala Asn Leu Met Trp Ala Ser Ser Leu Ala Ile Asn Gly Leu Leu
245 250 255

Thr Tyr Gly Lys Asp Thr Asn Trp Ser Val His Leu Met Glu His Glu
260 265 270

Leu Ser Ala Tyr Tyr Asp Ile Thr His Gly Val Gly Leu Ala Ile Leu
275 280 285

20130927_CL5646WOPCT_ST25. txt

Thr Pro Asn Trp Met Glu Tyr Ile Leu Asn Asn Asp Thr Val Tyr Lys
290 295 300

Phe Val Glu Tyr Gly Val Asn Val Trp Gly Ile Asp Lys Glu Lys Asn
305 310 315 320

His Tyr Asp Ile Ala His Gln Ala Ile Gln Lys Thr Arg Asp Tyr Phe
325 330 335

Val Asn Val Leu Gly Leu Pro Ser Arg Leu Arg Asp Val Gly Ile Glu
340 345 350

Glu Glu Lys Leu Asp Ile Met Ala Lys Glu Ser Val Lys Leu Thr Gly
355 360 365

Gly Thr Ile Gly Asn Leu Arg Pro Val Asn Ala Ser Glu Val Leu Gln
370 375 380

Ile Phe Lys Lys Ser Val
385 390

<210> 44
<211> 1173
<212> DNA
<213> Clostridium acetobutylicum

<400> 44
gtggttgatt tcgaatattc aataccaact agaatttttt tcggtaaaga taagataaat 60
gtacttggaa gagagcttaa aaaatatggt tctaaagtgc ttatagttaa tggtggagga 120
agtataaaga gaaatggaat atatgataaa gctgtaagta tacttgaaaa aaacagtatt 180
aaattttatg aacttgcagg agtagagcca aatccaagag taactacagt tgaaaaagga 240
gttaaaatat gtagagaaaa tggagttgaa gtagtactag ctataggtgg aggaagtgca 300
atagattgcg caaaggttat agcagcagca tgtgaatatg atggaaatcc atgggatatt 360
gtgtagatg gctcaaaaat aaaaaggtg cttcctatag ctagtatatt aaccattgct 420
gcaacaggat cagaaatgga tacgtgggca gtaataaata atatggatac aaacgaaaaa 480
ctaattgagg cacatccaga tatggctcct aagttttcta tattagatcc aacgtatacg 540
tataccgtac ctaccaatca aacagcagca ggaacagctg atattatgag tcatatattt 600
gaggtgtatt ttagtaatac aaaaacagca tatttgcagg atagaatggc agaagcgtaa 660
ttaagaactt gtattaaata tggaggaata gctcttgaga agccggatga ttatgaggca 720
agagccaatc taatgtgggc ttcaagtctt gcgataaatg gacttttaac atatggtaaa 780
gacactaatt ggagtgtaca cttaatggaa catgaattaa gtgcttatta cgacataaca 840
cacggcgtag ggcttgcaat tttacacct aattggatgg agtatatttt aaataatgat 900
acagtgtaca agtttgttga atatggtgta aatgtttggg gaatagacaa agaaaaaat 960
cactatgaca tagcacatca agcaatacaa aaaacaagag attactttgt aaatgtacta 1020

ggtttaccat ctagactgag agatgttgga attgaagaag aaaaattgga cataatggca 1080
aaggaatcag taaagcttac aggaggaacc ataggaaacc taagaccagt aaacgcctcc 1140
gaagtcctac aaatattcaa aaaatctgtg taa 1173

<210> 45
<211> 330
<212> PRT
<213> *Bacillus subtilis*

<400> 45

Met Ser Thr Asn Arg His Gln Ala Leu Gly Leu Thr Asp Gln Glu Ala
1 5 10 15

Val Asp Met Tyr Arg Thr Met Leu Leu Ala Arg Lys Ile Asp Glu Arg
20 25 30

Met Trp Leu Leu Asn Arg Ser Gly Lys Ile Pro Phe Val Ile Ser Cys
35 40 45

Gln Gly Gln Glu Ala Ala Gln Val Gly Ala Ala Phe Ala Leu Asp Arg
50 55 60

Glu Met Asp Tyr Val Leu Pro Tyr Tyr Arg Asp Met Gly Val Val Leu
65 70 75 80

Ala Phe Gly Met Thr Ala Lys Asp Leu Met Met Ser Gly Phe Ala Lys
85 90 95

Ala Ala Asp Pro Asn Ser Gly Gly Arg Gln Met Pro Gly His Phe Gly
100 105 110

Gln Lys Lys Asn Arg Ile Val Thr Gly Ser Ser Pro Val Thr Thr Gln
115 120 125

Val Pro His Ala Val Gly Ile Ala Leu Ala Gly Arg Met Glu Lys Lys
130 135 140

Asp Ile Ala Ala Phe Val Thr Phe Gly Glu Gly Ser Ser Asn Gln Gly
145 150 155 160

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

Ile Phe Met Cys Glu Asn Asn Lys Tyr Ala Ile Ser Val Pro Tyr Asp
180 185 190

Lys Gln Val Ala Cys Glu Asn Ile Ser Asp Arg Ala Ile Gly Tyr Gly
195 200 205

Met Pro Gly Val Thr Val Asn Gly Asn Asp Pro Leu Glu Val Tyr Gln

210

215

220

Ala Val Lys Glu Ala Arg Glu Arg Ala Arg Arg Gly Glu Gly Pro Thr
 225 230 235 240

Leu Ile Glu Thr Ile Ser Tyr Arg Leu Thr Pro His Ser Ser Asp Asp
 245 250 255

Asp Asp Ser Ser Tyr Arg Gly Arg Glu Glu Val Glu Glu Ala Lys Lys
 260 265 270

Ser Asp Pro Leu Leu Thr Tyr Gln Ala Tyr Leu Lys Glu Thr Gly Leu
 275 280 285

Leu Ser Asp Glu Ile Glu Gln Thr Met Leu Asp Glu Ile Met Ala Ile
 290 295 300

Val Asn Glu Ala Thr Asp Glu Ala Glu Asn Ala Pro Tyr Ala Ala Pro
 305 310 315 320

Glu Ser Ala Leu Asp Tyr Val Tyr Ala Lys
 325 330

<210> 46

<211> 993

<212> DNA

<213> *Bacillus subtilis*

<400> 46

atgagtacaa accgacatca agcactaggg ctgactgatac aggaagccgt tgatatgtat 60

agaaccatgc tgtagcaag aaaaatcgat gaaagaatgt ggctgttaaa ccgttctggc 120

aaaattccat ttgtaatctc ttgtcaagga caggaagcag cacaggtagg agcggctttc 180

gcacttgacc gtgaaatgga ttatgtattg ccgtactaca gagacatggg tgcgtgctc 240

gcgtttggca tgacagcaaa ggacttaatg atgtccgggt ttgcaaaagc agcagatccg 300

aactcaggag gccgccagat gccgggacat ttcggacaaa agaaaaaccg cattgtgacg 360

ggatcatctc cggttacaac gcaagtgcg cagcagtcg gtattgcgct tgcgggacgt 420

atggagaaaa aggatatcgc agcctttgtt acattcgggg aagggtcttc aaaccaaggc 480

gatttccatg aaggggcaaa ctttgccgct gtccataagc tgccggttat tttcatgtgt 540

gaaaacaaca aatacgcaat ctcatgcct tacgataagc aagtcgcatg tgagaacatt 600

tccgaccgtg ccataggcta tgggatgcct ggcgtaactg tgaatggaaa tgatccgctg 660

gaagtttatc aagcgggttaa agaagcacgc gaaagggcac gcagaggaga agggccgaca 720

ttaattgaaa cgatttctta ccgccttaca ccacattcca gtgatgacga tgacagcagc 780

tacagaggcc gtgaagaagt agaggaagcg aaaaaaagt atcccctgct tacttatcaa 840

gcttacttaa aggaaacagg cctgctgtcc gatgagatag acaaaccat gctggatgaa 900

attatggcaa tcgtaaata agcgacggat gaagcggaga acgcccata tgcagctcct 960

gagtcagcgc ttgattatgt ttatgcgaag tag

993

<210> 47
 <211> 327
 <212> PRT
 <213> Bacillus subtilis

<400> 47

Met Ser Val Met Ser Tyr Ile Asp Ala Ile Asn Leu Ala Met Lys Glu
 1 5 10 15

Glu Met Glu Arg Asp Ser Arg Val Phe Val Leu Gly Glu Asp Val Gly
 20 25 30

Arg Lys Gly Gly Val Phe Lys Ala Thr Ala Gly Leu Tyr Glu Gln Phe
 35 40 45

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

Gly Val Gly Ile Gly Ala Ala Met Tyr Gly Met Arg Pro Ile Ala Glu
 65 70 75 80

Met Gln Phe Ala Asp Phe Ile Met Pro Ala Val Asn Gln Ile Ile Ser
 85 90 95

Glu Ala Ala Lys Ile Arg Tyr Arg Ser Asn Asn Asp Trp Ser Cys Pro
 100 105 110

Ile Val Val Arg Ala Pro Tyr Gly Gly Gly Val His Gly Ala Leu Tyr
 115 120 125

His Ser Gln Ser Val Glu Ala Ile Phe Ala Asn Gln Pro Gly Leu Lys
 130 135 140

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

Ala Val Arg Asp Glu Asp Pro Val Leu Phe Phe Glu His Lys Arg Ala
 165 170 175

Tyr Arg Leu Ile Lys Gly Glu Val Pro Ala Asp Asp Tyr Val Leu Pro
 180 185 190

Ile Gly Lys Ala Asp Val Lys Arg Glu Gly Asp Asp Ile Thr Val Ile
 195 200 205

Thr Tyr Gly Leu Cys Val His Phe Ala Leu Gln Ala Ala Glu Arg Leu
 210 215 220

Glu Lys Asp Gly Ile Ser Ala His Val Val Asp Leu Arg Thr Val Tyr

225 230 — 235 — 240

Pro Leu Asp Lys Glu Ala Ile Ile Glu Ala Ala Ser Lys Thr Gly Lys
245 250 255

Val Leu Leu Val Thr Glu Asp Thr Lys Glu Gly Ser Ile Met Ser Glu
260 265 270

Val Ala Ala Ile Ile Ser Glu His Cys Leu Phe Asp Leu Asp Ala Pro
275 280 285

Ile Lys Arg Leu Ala Gly Pro Asp Ile Pro Ala Met Pro Tyr Ala Pro
290 295 300

Thr Met Gl u Lys Tyr Phe Met Val Asn Pro Asp Lys Val Gl u Al a Al a
305 310 315 320

Met Arg Glu Leu Ala Glu Phe
325

<210>	48
<211>	984
<212>	DNA
<213>	Bacillus subtilis

<400>	48						
atgtcagtaa	tgtcatatat	tgatgcaatc	aattttggcg	tgaagaaga	aatggaacga		60
gatttctcgcg	ttttcgtcct	tggggaagat	gtaggaagaa	aaggcgggtgt	gtttaaagcg		120
acagcggggac	tctatgaaca	atttggggaa	gagcgcgtta	tggatacgcc	gcttgctgaa		180
tctgcaatcg	caggagtcgg	tatcggagcg	gcaatgtacg	gaatgagacc	gattgctgaa		240
atgcagtttg	ctgatttcat	tatgccggca	gtcaaccaa	ttatttctga	agcggctaaa		300
atccgctacc	gcagcaaca	tgactggagc	tgtccgattg	tcgtcagagc	gccatacggc		360
ggaggcgtgc	acggagccct	gtatcattct	caatcagtcg	aagcaatttt	cgccaaccag		420
cccggactga	aaattgtcat	gccatcaaca	ccatatgacg	cgaaagggct	cttaaaagcc		480
gcagttcgtg	acgaagacc	cgtgctgttt	tttgagcaca	agcgggcata	ccgtctgata		540
aagggcgagg	ttccggctga	tgattatgtc	ctgccaatcg	gcaaggcgga	cgtaaaaagg		600
gaaggcgacg	acatcacagt	gatcacatac	ggcctgtgtg	tccacttcgc	cttacaagct		660
gcagaacgtc	tcgaaaaaga	tggcatttca	gcgcatgtgg	tggatttaag	aacagtttac		720
ccgcttgata	aagaagccat	catcgaagct	gcgtccaaaa	ctggaaaggt	tcttttggtc		780
acagaagata	caaaagaagg	cagcatcatg	agcgaagtag	ccgcaattat	atccgagcat		840
tgtctgttcg	acttagacgc	gccgatcaaa	cggcttgacg	gtcctgatat	tccggctatg		900
ccttatgcmc	cgacaatgga	aaaatacttt	atggtcaacc	ctgataaagt	ggaagcggcg		960
atgagagaat	tagcggagtt	ttaa					984

<210> 49
 <211> 424
 <212> PRT
 <213> Bacillus subtilis

<400> 49

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

Glu Gly Thr Ile Ser Lys Trp Leu Val Ala Pro Gly Asp Lys Val Asn
 20 25 30

Lys Tyr Asp Pro Ile Ala Glu Val Met Thr Asp Lys Val Asn Ala Glu
 35 40 45

Val Pro Ser Ser Phe Thr Gly Thr Ile Thr Glu Leu Val Gly Glu Glu
 50 55 60

Gly Gln Thr Leu Gln Val Gly Glu Met Ile Cys Lys Ile Glu Thr Glu
 65 70 75 80

Gly Ala Asn Pro Ala Glu Gln Lys Gln Glu Gln Pro Ala Ala Ser Glu
 85 90 95

Ala Ala Glu Asn Pro Val Ala Lys Ser Ala Gly Ala Ala Asp Gln Pro
 100 105 110

Asn Lys Lys Arg Tyr Ser Pro Ala Val Leu Arg Leu Ala Gly Glu His
 115 120 125

Gly Ile Asp Leu Asp Gln Val Thr Gly Thr Gly Ala Gly Gly Arg Ile
 130 135 140

Thr Arg Lys Asp Ile Gln Arg Leu Ile Glu Thr Gly Gly Val Gln Glu
 145 150 155 160

Gln Asn Pro Glu Glu Leu Lys Thr Ala Ala Pro Ala Pro Lys Ser Ala
 165 170 175

Ser Lys Pro Glu Pro Lys Glu Glu Thr Ser Tyr Pro Ala Ser Ala Ala
 180 185 190

Gly Asp Lys Glu Ile Pro Val Thr Gly Val Arg Lys Ala Ile Ala Ser
 195 200 205

Asn Met Lys Arg Ser Lys Thr Glu Ile Pro His Ala Trp Thr Met Met
 210 215 220

Glu Val Asp Val Thr Asn Met Val Ala Tyr Arg Asn Ser Ile Lys Asp
 225 230 235 240

Ser Phe Lys Lys Thr Glu Gly Phe Asn Leu Thr Phe Phe Ala Phe Phe
 Page 66

245

250

255

Val Lys Ala Val Ala Gln Ala Leu Lys Glu Phe Pro Gln Met Asn Ser
 260 265 270

Met Trp Ala Gly Asp Lys Ile Ile Gln Lys Lys Asp Ile Asn Ile Ser
 275 280 285

Ile Ala Val Ala Thr Glu Asp Ser Leu Phe Val Pro Val Ile Lys Asn
 290 295 300

Ala Asp Glu Lys Thr Ile Lys Gly Ile Ala Lys Asp Ile Thr Gly Leu
 305 310 315 320

Ala Lys Lys Val Arg Asp Gly Lys Leu Thr Ala Asp Asp Met Gln Gly
 325 330 335

Gly Thr Phe Thr Val Asn Asn Thr Gly Ser Phe Gly Ser Val Gln Ser
 340 345 350

Met Gly Ile Ile Asn Tyr Pro Gln Ala Ala Ile Leu Gln Val Glu Ser
 355 360 365

Ile Val Lys Arg Pro Val Val Met Asp Asn Gly Met Ile Ala Val Arg
 370 375 380

Asp Met Val Asn Leu Cys Leu Ser Leu Asp His Arg Val Leu Asp Gly
 385 390 395 400

Leu Val Cys Gly Arg Phe Leu Gly Arg Val Lys Gln Ile Leu Glu Ser
 405 410 415

Ile Asp Glu Lys Thr Ser Val Tyr
 420

<210> 50
 <211> 1275
 <212> DNA
 <213> Bacillus subtilis

<400> 50
 atggcaattg aacaaatgac gatgccgcag cttggagaaa gcgtaacaga ggggacgatc 60
 agcaaattggc ttgtcgcccc cggtgataaa gtgaacaaat acgatccgat cgcggaagtc 120
 atgacagata aggtaaatgc agaggttccg tcttctttta ctggtacgat aacagagctt 180
 gtgggagaag aaggccaaac cctgcaagtc ggagaaatga ttgcaaaat tgaaacagaa 240
 ggcgcgaatc cggctgaaca aaaacaagaa cagccagcag catcagaagc cgctgagaac 300
 cctgttgcaa aaagtgctgg agcagccgat cagcccaata aaaagcgcta ctcgccagct 360
 gttctccgtt tggccggaga gcacggcatt gacctcgatc aagtgcagg aactggtgcc 420
 ggcgggcgca tcacacgaaa agatattcag cgcttaattg aaacaggcgg cgtgcaagaa 480

```

cagaatcctg aggagctgaa aacagcagct cctgcaccga agtctgcatc aaaacctgag      540
ccaaaagaag agacgtcata tcctgctgtc gcagccggtg ataaagaaat ccctgtcaca      600
ggtgtaagaa aagcaattgc ttccaatatg aagcgaagca aaacagaaat tccgcatgct      660
tggacgatga tggaagtcga cgtcacaaat atggttgcac atcgcaacag tataaaagat      720
tcttttaaga agacagaagg ctttaattta acgttcttcg ctttttttgt aaaagcggtc      780
gctcaggcgt taaaagaatt cccgcaaatt aatagcatgt gggcggggga caaaattatt      840
cagaaaaagg atatcaatat ttcaattgca gttgccacag aggattcttt atttgttccg      900
gtgattaaaa acgctgatga aaaaacaatt aaaggcattg cgaaagacat taccggccta      960
gctaaaaaag taagagacgg aaaactcact gcagatgaca tgcagggagg cacgtttacc     1020
gtcaacaaca caggttcgtt cgggtctgtt cagtcgatgg gcattatcaa ctaccctcag     1080
gctgcgattc ttcaagtaga atccatcgtc aaacgcccggt ttgtcatgga caatggcatg     1140
attgctgtca gagacatggt taatctgtgc ctgtcattag atcacagagt gcttgacggt     1200
ctcgtgtgcg gacgattcct cggacgagtg aaacaaattt tagaatcgat tgacgagaag     1260
acatctgttt actaa                                                         1275

```

```

<210> 51
<211> 474
<212> PRT
<213> Bacillus subtilis
<400> 51

```

```

Met Ala Thr Glu Tyr Asp Val Val Ile Leu Gly Gly Gly Thr Gly Gly
1          5          10
Tyr Val Ala Ala Ile Arg Ala Ala Gln Leu Gly Leu Lys Thr Ala Val
          20          25          30
Val Glu Lys Glu Lys Leu Gly Gly Thr Cys Leu His Lys Gly Cys Ile
          35          40          45
Pro Ser Lys Ala Leu Leu Arg Ser Ala Glu Val Tyr Arg Thr Ala Arg
          50          55          60
Glu Ala Asp Gln Phe Gly Val Glu Thr Ala Gly Val Ser Leu Asn Phe
65          70          75          80
Glu Lys Val Gln Gln Arg Lys Gln Ala Val Val Asp Lys Leu Ala Ala
          85          90          95
Gly Val Asn His Leu Met Lys Lys Gly Lys Ile Asp Val Tyr Thr Gly
          100          105          110
Tyr Gly Arg Ile Leu Gly Pro Ser Ile Phe Ser Pro Leu Pro Gly Thr
          115          120          125

```

20130927_CL5646WOPCT_ST25. txt

I l e S e r V a l G l u A r g G l y A s n G l y G l u G l u A s n A s p M e t L e u I l e P r o
130 135 140

L y s G l n V a l I l e I l e A l a T h r G l y S e r A r g P r o A r g M e t L e u P r o G l y
145 150 155 160

L e u G l u V a l A s p G l y L y s S e r V a l L e u T h r S e r A s p G l u A l a L e u G l n
165 170 175

M e t G l u G l u L e u P r o G l n S e r I l e I l e I l e V a l G l y G l y G l y V a l I l e
180 185 190

G l y I l e G l u T r p A l a S e r M e t L e u H i s A s p P h e G l y V a l L y s V a l T h r
195 200 205

V a l I l e G l u T y r A l a A s p A r g I l e L e u P r o T h r G l u A s p L e u G l u I l e
210 215 220

S e r L y s G l u M e t G l u S e r L e u L e u L y s L y s L y s G l y I l e G l n P h e I l e
225 230 235 240

T h r G l y A l a L y s V a l L e u P r o A s p T h r M e t T h r L y s T h r S e r A s p A s p
245 250 255

I l e S e r I l e G l n A l a G l u L y s A s p G l y G l u T h r V a l T h r T y r S e r A l a
260 265 270

G l u L y s M e t L e u V a l S e r I l e G l y A r g G l n A l a A s n I l e G l u G l y I l e
275 280 285

G l y L e u G l u A s n T h r A s p I l e V a l T h r G l u A s n G l y M e t I l e S e r V a l
290 295 300

A s n G l u S e r C y s G l n T h r L y s G l u S e r H i s I l e T y r A l a I l e G l y A s p
305 310 315 320

V a l I l e G l y G l y L e u G l n L e u A l a H i s V a l A l a S e r H i s G l u G l y I l e
325 330 335

I l e A l a V a l G l u H i s P h e A l a G l y L e u A s n P r o H i s P r o L e u A s p P r o
340 345 350

T h r L e u V a l P r o L y s C y s I l e T y r S e r S e r P r o G l u A l a A l a S e r V a l
355 360 365

G l y L e u T h r G l u A s p G l u A l a L y s A l a A s n G l y H i s A s n V a l L y s I l e
370 375 380

G l y L y s P h e P r o P h e M e t A l a I l e G l y L y s A l a L e u V a l T y r G l y G l u
385 390 395 400

20130927_CL5646WOPCT_ST25. txt

Ser Asp Gly Phe Val Lys Ile Val Ala Asp Arg Asp Thr Asp Asp Ile
405 410 415

Leu Gly Val His Met Ile Gly Pro His Val Thr Asp Met Ile Ser Glu
420 425 430

Ala Gly Leu Ala Lys Val Leu Asp Ala Thr Pro Trp Glu Val Gly Gln
435 440 445

Thr Ile His Pro His Pro Thr Leu Ser Glu Ala Ile Gly Glu Ala Ala
450 455 460

Leu Ala Ala Asp Gly Lys Ala Ile His Phe
465 470

<210> 52
<211> 1425
<212> DNA
<213> Bacillus subtilis

<400> 52
atggcaactg agtatgacgt agtcattctg ggcggcggta ccggcgggta tgttgcggcc 60
atcagagccg ctcagctcgg cttaaaaaca gccgttgttg aaaaggaaaa actcggggga 120
acatgtctgc ataaaggctg tatcccagat aaagcgtgc ttagaagcgc agaggtatac 180
cggacagctc gtgaagccga tcaattcggg gtggaaacgg ctggcgtgtc cctcaacttt 240
gaaaaagtgc agcagcgtaa gcaagccgtt gttgataagc ttgcagcggg tgtaaatacat 300
ttaatgaaaa aaggaaaaat tgacgtgtac accggatatg gacgtatcct tggaccgtca 360
atcttctctc cgctgccggg aacaatttct gttgagcggg gaaatggcga agaaaatgac 420
atgctgatcc cgaaacaagt gatcattgca acaggatcaa gaccgagaat gcttccgggt 480
cttgaagtgg acggttaagtc tgtactgact tcagatgagg cgctccaaat ggaggagctg 540
ccacagtcaa tcatcattgt cggcggaggg gttatcggtg tcgaatgggc gtctatgctt 600
catgattttg gcgttaaggt aacggttatt gaatacgcgg atcgcatatt gccgactgaa 660
gatctagaga tttcaaaaaga aatggaaagt cttcttaaga aaaaaggcat ccagttcata 720
acagggggcaa aagtgtctgc tgacacaatg acaaaaacat cagacgatat cagcatataa 780
gcggaaaaag acggagaaac cgttacctat tctgctgaga aaatgcttgt ttccatcggc 840
agacaggcaa atatcgaagg catcggccta gagaacaccg atattgttac tgaaaatggc 900
atgatttcag tcaatgaaag ctgccaaacg aaggaatctc atatttatgc aatcggagac 960
gtaatcggtg gcctgcagtt agctcacgtt gcttcacatg aggaattat tgctgttgag 1020
cattttgcag gtctcaatcc gcatccgctt gatccgacgc ttgtgccgaa gtgcatttac 1080
tcaagccctg aagctgccag tgtcggctta accgaagacg aagcaaaggc gaacgggcat 1140
aatgtcaaaa tcggcaagtt cccatttatg gcgattggaa aagcgcttgt atacggtgaa 1200
agcgacgggt ttgtcaaaat cgtggctgac cgagatacag atgatattct cggcgttcat 1260

atgattggcc cgcatgtcac cgacatgatt tctgaagcgg gtcttgccaa agtgctggac 1320
gcaacaccgt gggaggtcgg gcaaacgatt caccgcgcatc caacgctttc tgaagcaatt 1380
ggagaagctg cgcttgccgc agatggcaaa gccattcatt ttttaa 1425

<210> 53
<211> 410
<212> PRT
<213> Pseudomonas putida

<400> 53

Met Asn Glu Tyr Ala Pro Leu Arg Leu His Val Pro Glu Pro Thr Gly
1 5 10 15

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

Gly Gln Ala Arg Lys Pro Pro Val Asp Val Asp Ala Ala Asp Thr Ala
35 40 45

Asp Leu Ser Tyr Ser Leu Val Arg Val Leu Asp Glu Gln Gly Asp Ala
50 55 60

Gln Gly Pro Trp Ala Glu Asp Ile Asp Pro Gln Ile Leu Arg Gln Gly
65 70 75 80

Met Arg Ala Met Leu Lys Thr Arg Ile Phe Asp Ser Arg Met Val Val
85 90 95

Ala Gln Arg Gln Lys Lys Met Ser Phe Tyr Met Gln Ser Leu Gly Glu
100 105 110

Glu Ala Ile Gly Ser Gly Gln Ala Leu Ala Leu Asn Arg Thr Asp Met
115 120 125

Cys Phe Pro Thr Tyr Arg Gln Gln Ser Ile Leu Met Ala Arg Asp Val
130 135 140

Ser Leu Val Glu Met Ile Cys Gln Leu Leu Ser Asn Glu Arg Asp Pro
145 150 155 160

Leu Lys Gly Arg Gln Leu Pro Ile Met Tyr Ser Val Arg Glu Ala Gly
165 170 175

Phe Phe Thr Ile Ser Gly Asn Leu Ala Thr Gln Phe Val Gln Ala Val
180 185 190

Gly Trp Ala Met Ala Ser Ala Ile Lys Gly Asp Thr Lys Ile Ala Ser
195 200 205

Ala Trp Ile Gly Asp Gly Ala Thr Ala Glu Ser Asp Phe His Thr Ala
Page 71

210

215

220

Leu Thr Phe Ala His Val Tyr Arg Ala Pro Val Ile Leu Asn Val Val
225 230 235 240

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

Ser Thr Thr Phe Ala Gly Arg Gly Val Gly Cys Gly Ile Ala Ser Leu
260 265 270

Arg Val Asp Gly Asn Asp Phe Val Ala Val Tyr Ala Ala Ser Arg Trp
275 280 285

Ala Ala Glu Arg Ala Arg Arg Gly Leu Gly Pro Ser Leu Ile Glu Trp
290 295 300

Val Thr Tyr Arg Ala Gly Pro His Ser Thr Ser Asp Asp Pro Ser Lys
305 310 315 320

Tyr Arg Pro Ala Asp Asp Trp Ser His Phe Pro Leu Gly Asp Pro Ile
325 330 335

Ala Arg Leu Lys Gln His Leu Ile Lys Ile Gly His Trp Ser Glu Glu
340 345 350

Glu His Gln Ala Thr Thr Ala Glu Phe Glu Ala Ala Val Ile Ala Ala
355 360 365

Gln Lys Glu Ala Glu Gln Tyr Gly Thr Leu Ala Asn Gly His Ile Pro
370 375 380

Ser Ala Ala Ser Met Phe Glu Asp Val Tyr Lys Glu Met Pro Asp His
385 390 395 400

Leu Arg Arg Gln Arg Gln Glu Leu Gly Val
405 410

<210> 54

<211> 6643

<212> DNA

<213> Pseudomonas putida

<400> 54

gcatgcctgc aggccgccga tgaaatggtg gaaggtatcg gtaggctggc cctgctcatc 60

gctgaacacg ttacgcccgc tgccggtatc gaccaggctc tggatgaatat gcatggaact 120

gccaggcgtg cgcgccagcg gtttggccat gcacaccacg gtcagcccgt gcttgagtgc 180

cacttccttg agcagggtgtt tgaacaggaa ggtctggtcg gccagcagca gcgggtcgcc 240

atgtagcaag ttgatctcga actggctgac gccatttcg tgcatgaagg tgtcgcgcgg 300

caggccgagc gcggccatgc actggtacac ctattgaag aacgggcgca ggccgttggt 360

20130927_CL5646WOPCT_ST25. txt

ggaactgaca	ctgaacgccg	aatggcccag	ctcgcggcgg	ccgtcgggtgc	ccagcgggtgg	420
ctggaacggc	tgctgcgggt	cactgtttgg	ggcaaacacg	aagaactcaa	gctcgggtcgc	480
cactaccggt	gccagaccca	acgctgcgta	gcgggcgatc	acggccttca	gctggccccg	540
ggtggacagt	gccgagggcc	ggccatccag	ttcattggca	tcgcagatgg	ccagggcgcg	600
accgtcatcg	ctccagggga	agcgatgaac	ctggctgggt	tccgctacca	acgccaggtc	660
gccgtcgtcg	cagccgtaga	atttcgccgg	cgggtagccg	cccatgatgc	attgcagcag	720
cacccacagg	gccatctgca	ggcggcgggc	ttcagaaaag	ccttcggcgg	tcatacctt	780
gccgcgtggg	acgccgttga	ggtcgggggt	gacgcattcg	atttcatcga	tgccctggag	840
ctgagcgatg	ctcatgacgc	ttgtccttgt	tgttgtaggc	tgacaacaac	ataggctggg	900
ggtgtttaaa	atatcaagca	gcctctcgaa	cgcctggggc	ctcttctatt	cgcgcaagggt	960
catgccattg	gccggcaacg	gcaaggctgt	cttgtagcgc	acctgtttca	aggcaaaact	1020
cgagcggata	ttcgccacac	ccggcaaccg	ggtcaggtaa	tcgagaaacc	gctccagcgc	1080
ctggatactc	ggcagcagta	cccgaacag	gtagtccggg	tcgcccgtca	tcaggtagca	1140
ctccatcacc	tcggggccgtt	cggcaatttc	ttcctcgaag	cgggtgcagcg	actgctctac	1200
ctgtttttcc	aggctgacat	ggatgaacac	attcacatcc	agccccaacg	cctcggggcga	1260
caacaaggtc	acctgctggc	ggatcacccc	cagttcttcc	atggcccgca	cccggttgaa	1320
acagggcggtg	ggcgacaggt	tgaccgagcg	tgccagctcg	gcgttggtga	tgcgggcggtt	1380
ttcctgcagg	ctgttgagaa	tgccgatatc	ggtacgatcg	agtttgcgca	tgagacaaaa	1440
tcaccggttt	tttgtgttta	tgcggaatgt	ttatctgccc	cgctcggcaa	aggcaatcaa	1500
cttgagagaa	aaattctcct	gccggaccac	taagatgtag	gggacgctga	cttaccagtc	1560
acaagccggt	actcagcggc	ggccgcttca	gagctcacia	aaacaaatac	ccgagcgagc	1620
gtaaaaagca	tgaacgagta	cgccccctg	cgtttgcatg	tgcccagagcc	caccggccgg	1680
ccagggtgcc	agaccgattt	ttcctacctg	cgcctgaacg	atgcagggtca	agcccgtaaa	1740
ccccctgtcg	atgtcgacgc	tgccgacacc	gccgacctgt	cctacagcct	ggtccgcgtg	1800
ctcgacgagc	aaggcgacgc	ccaaggcccc	tggtgtgaag	acatcgaccc	gcagatcctg	1860
cgccaaggca	tgcgcgccat	gctcaagacg	cggatcttcg	acagccgcat	ggtggttgcc	1920
cagcgccaga	agaagatgtc	cttctacatg	cagagcctgg	gcgaagaagc	catcggcagc	1980
ggccaggcgc	tggtcgcttaa	ccgcaccgac	atgtgcttcc	ccacctaccg	tcagcaaagc	2040
atcctgatgg	cccgcgacgt	gtcgctggtg	gagatgatct	gccagttgct	gtccaacgaa	2100
cgcgaccccc	tcaagggccg	ccagctgccg	atcatgtact	cggtagcgga	ggccggcttc	2160
ttcaccatca	gcggcaacct	ggcgacccag	ttcgtgcagg	cggtcgggtg	ggccatggcc	2220
tcggcgatca	agggcgatac	caagattgcc	tcggcctgga	tcggcgacgg	cgccactgcc	2280
gaatcggact	tccacaccgc	cctcaccttt	gcccacgttt	accgcgcccc	ggtgatcctc	2340
aacgtggtca	acaaccagtg	ggccatctca	accttcagg	ccatcgccgg	tggtgagtcg	2400

20130927_CL5646WOPCT_ST25. txt

accaccttcg	ccggccgtgg	cgtgggctgc	ggcatcgctt	cgctgcgggt	ggacggcaac	2460
gacttcgtcg	ccgtttacgc	cgcttcgcgc	tgggctgccg	aacgtgcccg	ccgtggtttg	2520
ggcccagacc	tgatcgagtg	ggtcacctac	cgtgccggcc	cgcactcgac	ctcggacgac	2580
ccgtccaagt	accgccctgc	cgatgactgg	agccacttcc	cgctgggtga	cccgatcgcc	2640
cgcctgaagc	agcacctgat	caagatcggc	cactgggtccg	aagaagaaca	ccaggccacc	2700
acggccgagt	tcgaagcggc	cgtgattgct	gcgcaaaaag	aagccgagca	gtacggcacc	2760
ctggccaacg	gtcacatccc	gagcgccgcc	tcgatgttcg	aggacgtgta	caaggagatg	2820
cccgaccacc	tgcgccgcca	acgccaggaa	ctgggggttt	gagatgaacg	accacaacaa	2880
cagcatcaac	ccggaaaccg	ccatggccac	cactaccatg	accatgatcc	aggccctgcg	2940
ctcggccatg	gatgtcatgc	ttgagcgcg	cgacaatgtg	gtggtgtacg	gccaggacgt	3000
cggctacttc	ggcggcgtgt	tccgctgcac	cgaaggcctg	cagaccaagt	acggcaagtc	3060
ccgcgtgttc	gacgcgcccc	tctctgaaag	cggcatcgtc	ggcaccgccg	tgggcatggg	3120
tgcctacggc	ctgcgcccgg	tggtggaat	ccagttcgct	gactacttct	acccggcctc	3180
cgaccagatc	gtttctgaaa	tggcccgct	gcgctaccgt	tcggccggcg	agttcatcgc	3240
cccgtgacc	ctgcgtatgc	cctgcgggtg	cggtatctat	ggcggccaga	cacacagcca	3300
gagcccggaa	gcgatgttca	ctcagggtgtg	cggcctgcgc	accgtaatgc	catccaaccc	3360
gtacgacgcc	aaaggcctgc	tgattgcctc	gatcgaatgc	gacgaccggg	tgatcttctt	3420
ggagcccaag	cgcctgtaca	acggcccgtt	cgacggccac	catgaccgcc	cggttacgcc	3480
gtggtcgaaa	cacccgcaca	gcgccgtgcc	cgatggctac	tacaccgtgc	cactggacaa	3540
ggccgcatc	acccgccccg	gcaatgacgt	gagcgtgctc	acctatggca	ccaccgtgta	3600
cgtggcccag	gtggccggcg	aagaaagtgg	cgtggatgcc	gaagtgatcg	acctgcgcag	3660
cctgtggccg	ctagacctgg	acaccatcgt	cgagtcgggtg	aaaaagaccg	gccgttgctg	3720
ggtagtacac	gaggccaccc	gtacttgtgg	ctttggcgca	gaactggtgt	cgctggtgca	3780
ggagcactgc	ttccaccacc	tggaggcgcc	gatcgagcgc	gtcaccggtt	gggacacccc	3840
ctaccctcac	gcgcaggaat	gggcttactt	cccagggcct	tcgcgggtag	gtgcggcatt	3900
gaaaaaggtc	atggaggctt	gaatgggcac	gcacgtcatc	aagatgccgg	acattggcga	3960
aggcatcgcg	caggtcgaat	tggtggaatg	gttcgtcaag	gtgggcgaca	tcacgcgcga	4020
ggaccaagtg	gtagccgacg	tcatgaccga	caaggccacc	gtggaaatcc	cgtcgccggt	4080
cagcggcaag	gtgctggccc	tgggtggcca	gccaggtgaa	gtgatggcgg	tcggcagtga	4140
gctgatccgc	atcgaagtgg	aaggcagcgg	caaccatgtg	gatgtgccgc	aagccaagcc	4200
ggccgaagtg	cctgcggcac	cggtagccgc	taaacctgaa	ccacagaaaag	acgttaaacc	4260
ggcggcgtag	caggcgtagc	ccagccacga	ggcagcgccc	atcgtgccgc	gccagccggg	4320
cgacaagccg	ctggcctcgc	cggcggtgcg	caaacgcgcc	ctcgatgccg	gcacgaattt	4380
gcgttatgtg	cacggcagcg	gcccggccgg	gcgcatcctg	cacgaagacc	tcgacgcgtt	4440

20130927_CL5646WOPCT_ST25. txt

catgagcaaa	ccgcaaagcg	ctgccgggca	aacccccaat	ggctatgcca	ggcgcaccga	4500
cagcgagcag	gtgccggtga	tcggcctgcg	ccgcaagatc	gcccagcgca	tgcaggacgc	4560
caagcgccgg	gtcgcgcact	tcagctatgt	ggaagaaatc	gacgtcaccg	ccctggaagc	4620
cctgcgccag	cagctcaaca	gcaagcacgg	cgacagccgc	ggcaagctga	cactgctgcc	4680
gttcctggtg	cgcgccctgg	tcgtggcact	gcgtgacttc	ccgcagataa	acgccaccta	4740
cgatgacgaa	gcgcagatca	tcacccgcca	tggcgcggtg	catgtgggca	tcgccaccca	4800
aggtgacaac	ggcctgatgg	tacccgtgct	gcgccacgcc	gaagcgggca	gcctgtgggc	4860
caatgccggt	gagatttcac	gcctggccaa	cgctgcgcgc	aacaacaagg	ccagccgcga	4920
agagctgtcc	ggttcgacca	ttaccctgac	cagcctcggc	gccctgggcg	gcacgtcag	4980
cacgccggtg	gtcaacaccc	cgaagtggc	gatcgtcgg	gtcaaccgca	tggttgagcg	5040
gcccggtggtg	atcgacggcc	agatcgtcgt	gcgcaagatg	atgaacctgt	ccagctcgtt	5100
cgaccaccgc	gtggtcgatg	gcatggacgc	cgccctgttc	atccaggccg	tgcgtggcct	5160
gctcgaacaa	cccgcctgcc	tgttcgtgga	gtgagcatgc	aacagactat	ccagacaacc	5220
ctgttgatca	tcggcggcgg	ccctggcggc	tatgtggcgg	ccatccgcgc	cgggcaactg	5280
ggcatcccta	ccgtgctggt	ggaaggccag	gcgctgggcg	gtacctgcct	gaacatcggc	5340
tgcatccgt	ccaaggcgct	gatccatgtg	gccgagcagt	tccaccaggc	ctcgcgcttt	5400
accgaaccct	cgccgctggg	catcagcgtg	gcttcgccac	gcctggacat	cggccagagc	5460
gtggcctgga	aagacggcat	cgtcgatcgc	ctgaccactg	gtgtcgccgc	cctgctgaaa	5520
aagcacgggg	tgaagggtgt	gcacggctgg	gccaagggtg	ttgatggcaa	gcaggtcgag	5580
gtggatggcc	agcgcatcca	gtgcgagcac	ctgttgctgg	ccacgggctc	cagcagtgtc	5640
gaactgccga	tgctgccgtt	gggtgggccc	gtgatttcct	cgaccgaggc	cctggcaccg	5700
aaagccctgc	cgcaacacct	ggtggtggtg	ggcgggtggc	acatcggcct	ggagctgggt	5760
atgcctacc	gcaagctcgg	cgcgcaggtc	agcgtggtgg	aagcgcgcga	gcgcacacct	5820
ccgacttacg	acagcgaact	gaccgccccg	gtggccgagt	cgctgaaaaa	gctgggtatc	5880
gccctgcacc	ttggccacag	cgtcgaagg	tacgaaaatg	gctgcctgct	ggccaacgat	5940
ggcaagggcg	gacaactgcg	cctggaagcc	gaccgggtgc	tgggtggcgt	gggccgccgc	6000
ccacgcacca	agggttcaa	cctggaatgc	ctggacctga	agatgaatgg	tgccgcgatt	6060
gccatcgacg	agcgtgccca	gaccagcatg	cacaacgtct	gggccatcgg	cgacgtggcc	6120
ggcgaaccga	tgctggcgca	ccgggccatg	gccagggcg	agatggtggc	cgagatcatc	6180
gccggcaagg	cacgccgctt	cgaacccgct	gcgatagccg	ccgtgtgctt	caccgaccgc	6240
gaagtggctg	tggtcggcaa	gacgccggaa	caggccagtc	agcaaggcct	ggactgcac	6300
gtcgcgcagt	tcccgttcgc	cgccaacggc	cgggccatga	gcctggagtc	gaaaagcggt	6360
ttcgtgcgcg	tggtcgcgcg	gcgtgacaac	cacctgatcc	tgggctggca	agcggttggc	6420
gtggcgggtt	ccgagctgtc	cacggcggtt	gccagtcgc	tggagatggg	tgctgcctg	6480

gaggatgtgg ccggtacat ccatgcccac ccgaccctgg gtgaagcggg acaggaagcg 6540
gcactgctg ccctgggcca cgccctgcat atctgacact gaagcggccg aggccgattt 6600
ggcccgccgc gccgagaggc gctgcgggtc tttttatac ctg 6643

<210> 55
<211> 352
<212> PRT
<213> Pseudomonas putida

<400> 55

Met Asn Asp His Asn Asn Ser Ile Asn Pro Glu Thr Ala Met Ala Thr
1 5 10 15

Thr Thr Met Thr Met Ile Gln Ala Leu Arg Ser Ala Met Asp Val Met
20 25 30

Leu Glu Arg Asp Asp Asn Val Val Val Tyr Gly Gln Asp Val Gly Tyr
35 40 45

Phe Gly Gly Val Phe Arg Cys Thr Glu Gly Leu Gln Thr Lys Tyr Gly
50 55 60

Lys Ser Arg Val Phe Asp Ala Pro Ile Ser Glu Ser Gly Ile Val Gly
65 70 75 80

Thr Ala Val Gly Met Gly Ala Tyr Gly Leu Arg Pro Val Val Glu Ile
85 90 95

Gln Phe Ala Asp Tyr Phe Tyr Pro Ala Ser Asp Gln Ile Val Ser Glu
100 105 110

Met Ala Arg Leu Arg Tyr Arg Ser Ala Gly Glu Phe Ile Ala Pro Leu
115 120 125

Thr Leu Arg Met Pro Cys Gly Gly Gly Ile Tyr Gly Gly Gln Thr His
130 135 140

Ser Gln Ser Pro Glu Ala Met Phe Thr Gln Val Cys Gly Leu Arg Thr
145 150 155 160

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

Ile Glu Cys Asp Asp Pro Val Ile Phe Leu Glu Pro Lys Arg Leu Tyr
180 185 190

Asn Gly Pro Phe Asp Gly His His Asp Arg Pro Val Thr Pro Trp Ser
195 200 205

Lys His Pro His Ser Ala Val Pro Asp Gly Tyr Tyr Thr Val Pro Leu
Page 76

210

215

220

Asp Lys Ala Ala Ile Thr Arg Pro Gly Asn Asp Val Ser Val Leu Thr
 225 230 235 240

Tyr Gly Thr Thr Val Tyr Val Ala Gln Val Ala Ala Glu Glu Ser Gly
 245 250 255

Val Asp Ala Glu Val Ile Asp Leu Arg Ser Leu Trp Pro Leu Asp Leu
 260 265 270

Asp Thr Ile Val Glu Ser Val Lys Lys Thr Gly Arg Cys Val Val Val
 275 280 285

His Glu Ala Thr Arg Thr Cys Gly Phe Gly Ala Glu Leu Val Ser Leu
 290 295 300

Val Gln Glu His Cys Phe His His Leu Glu Ala Pro Ile Glu Arg Val
 305 310 315 320

Thr Gly Trp Asp Thr Pro Tyr Pro His Ala Gln Glu Trp Ala Tyr Phe
 325 330 335

Pro Gly Pro Ser Arg Val Gly Ala Ala Leu Lys Lys Val Met Glu Val
 340 345 350

<210> 56
 <211> 6643
 <212> DNA
 <213> Pseudomonas putida

<400> 56
 gcatgcctgc aggccgccga tgaaatggtg gaaggtatcg gtaggctggc cctgctcatc 60
 gctgaacacg ttacgcccgc tgccggtatc gaccaggctc tggatgaatat gcatggaact 120
 gccaggcgtg cgcgccagcg gtttgccat gcacaccacg gtcagcccgt gcttgagtgc 180
 cacttccttg agcaggtgtt tgaacaggaa ggtctggtcg gccagcagca gcgggtcgcc 240
 atgtagcaag ttgatctcga actggctgac gccatttcg tgcataagg tgcgcgcgg 300
 caggccgagc gcggccatgc actggtacac ctattgaag aacgggcgca ggccgttgtt 360
 ggaactgaca ctgaacgccg aatggcccag ctgcggcgcc ccgtcggtgc ccagcggtag 420
 ctggaacggc tgctgcgggt cactgttggg ggcaaacacg aagaactcaa gtcgggtcgc 480
 cactaccggt gccagacca acgctgcgta gcgggcgatc acggccttca gctggccccg 540
 ggtggacagt gccgagggcc ggccatccag ttattggca tcgcagatgg ccagggcgcg 600
 accgtcatcg ctccagggca agcagatgaac ctggctgggt tccgctacca acgccaggtc 660
 gccgtcgtcg cagccgtaga atttcgccg cggttagccg cccatgatgc attgcagcag 720
 caccacacgg gccatctgca ggcgggcgcc ttcagaaaag ccttcggcg tcatcacctt 780
 gccgcgtggg acgccgttga ggtcgggggt gacgcattcg atttcacga tgccctggag 840

20130927_CL5646WOPCT_ST25. txt

ctgagcgatg	ctcatgacgc	ttgtccttgt	tgttgtaggc	tgacaacaac	ataggctggg	900
ggtgttttaa	atatcaagca	gcctctcgaa	cgcctggggc	ctcttctatt	cgcgcaaggt	960
catgccattg	gccggcaacg	gcaaggctgt	cttgtagcgc	acctgtttca	aggcaaaact	1020
cgagcggata	ttcgccacac	ccggcaaccg	ggtcaggtaa	tcgagaaacc	gctccagcgc	1080
ctggatactc	ggcagcagta	cccgaacag	gtagtccggg	tcgcccgtca	tcaggtagca	1140
ctccatcacc	tcggggccgtt	cggcaatttc	ttcctcgaag	cggtgcagcg	actgctctac	1200
ctgtttttcc	aggctgacat	ggatgaacac	attcacatcc	agccccaacg	cctcgggcga	1260
caacaaggtc	acctgctggc	ggatcacccc	cagttcttcc	atggcccgcga	cccggttgaa	1320
acagggcgctg	ggcgacaggt	tgaccgagcg	tgccagctcg	gcgttggtga	tgcgggcggtt	1380
ttcctgcagg	ctgttgagaa	tgccgatatc	ggtacgatcg	agtttgcgca	tgagacaaaa	1440
tcaccggttt	tttgtgttta	tgcggaatgt	ttatctgccc	cgctcggcaa	aggcaatcaa	1500
cttgagagaa	aaattctcct	gccggaccac	taagatgtag	gggacgctga	cttaccagtc	1560
acaagccggt	actcagcggc	ggccgcttca	gagctcacia	aaacaaatac	ccgagcgagc	1620
gtaaaaagca	tgaacgagta	cgccccctg	cgtttgcatg	tgcccagagcc	caccggccgg	1680
ccaggctgcc	agaccgattt	ttcctacctg	cgcctgaacg	atgcagggtca	agcccgtaaa	1740
ccccctgtcg	atgtcgacgc	tgccgacacc	gccgacctgt	cctacagcct	ggtccgcgtg	1800
ctcgacgagc	aaggcgacgc	ccaaggcccc	tgggctgaag	acatcgaccc	gcagatcctg	1860
cgccaaggca	tgcgcgccat	gctcaagacg	cggatcttcg	acagccgcat	ggtgggttgc	1920
cagcgccaga	agaagatgtc	cttctacatg	cagagcctgg	gcgaagaagc	catcggcagc	1980
ggccaggcgc	tggcgcttaa	ccgcaccgac	atgtgcttcc	ccacctaccg	tcagcaaagc	2040
atcctgatgg	cccgcgacgt	gtcgctggtg	gagatgatct	gccagttgct	gtccaacgaa	2100
cgcgaccccc	tcaagggccg	ccagctgccg	atcatgtact	cggtacgcga	ggccggcttc	2160
ttcaccatca	gcggcaacct	ggcgaccag	ttcgtgcagg	cggtcggctg	ggccatggcc	2220
tcggcgatca	agggcgatac	caagattgcc	tcggcctgga	tcggcgacgg	cgccactgcc	2280
gaatcggact	tccacaccgc	cctcaccttt	gccacgttt	accgcgcccc	ggtgatcctc	2340
aacgtggtca	acaaccagtg	ggccatctca	accttcagg	ccatcgccgg	tggcgagtcg	2400
accaccttcg	ccggccgtgg	cgtgggctgc	ggcatcgctt	cgctgcgggt	ggacggcaac	2460
gacttcgtcg	ccgtttacgc	cgcttcgcgc	tgggctgccg	aacgtgcccc	ccgtggtttg	2520
ggcccagagc	tgatcgagtg	ggtcacctac	cgtgccggcc	cgactcgac	ctcggacgac	2580
ccgtccaagt	accgccctgc	cgatgactgg	agccacttcc	cgctgggtga	cccgatcgcc	2640
cgcctgaagc	agcacctgat	caagatcggc	cactgggtccg	aagaagaaca	ccaggccacc	2700
acggccgagt	tcgaagcggc	cgtgattgct	gcgaaaaaag	aagccgagca	gtacggcacc	2760
ctggccaacg	gtcacatccc	gagcgccgcc	tcgatgttcg	aggacgtgta	caaggagatg	2820
cccgaccacc	tgcgccgcca	acgccaggaa	ctgggggttt	gagatgaacg	accacaacaa	2880

20130927_CL5646WOPCT_ST25. txt

cagcatcaac	ccggaaccg	ccatggccac	cactaccatg	accatgatcc	aggccctgcg	2940
ctcggccatg	gatgtcatgc	ttgagcgcg	cgacaatgtg	gtggtgtacg	gccaggacgt	3000
cggctacttc	ggcggcgtgt	tccgctgcac	cgaaggcctg	cagaccaagt	acggcaagtc	3060
ccgcgtgttc	gacgcgcca	tctctgaaag	cggcatcgtc	ggcaccgccg	tgggcatggg	3120
tgcctacggc	ctgcgcccgg	tgggtggaaat	ccagttcgct	gactacttct	acccggcctc	3180
cgaccagatc	gtttctgaaa	tggcccgccct	gcgctaccgt	tcggccggcg	agttcatcgc	3240
cccgtgacc	ctgcgtatgc	cctgcggtgg	cggtatctat	ggcggccaga	cacacagcca	3300
gagcccggaa	gcgatgttca	ctcaggtgtg	cggcctgcgc	accgtaatgc	catccaaccc	3360
gtacgacgcc	aaaggcctgc	tgattgcctc	gatcgaatgc	gacgaccggg	tgatcttctt	3420
ggagcccaag	cgctgtaca	acggcccgtt	cgacggccac	catgaccgcc	cggttacgcc	3480
gtggtcgaaa	caccgcaca	gcgccgtgcc	cgatggctac	tacaccgtgc	cactggacaa	3540
ggccgccatc	acccgccccg	gcaatgacgt	gagcgtgctc	acctatggca	ccaccgtgta	3600
cgtggcccag	gtggccgccg	aagaaagtgg	cgtggatgcc	gaagtgatcg	acctgcgcag	3660
cctgtggccg	ctagacctgg	acaccatcgt	cgagtcggtg	aaaaagaccg	gccgttgctg	3720
ggtagtacac	gaggccaccc	gtacttgtgg	ctttggcgca	gaactggtgt	cgctggtgca	3780
ggagcactgc	ttccaccacc	tggaggcgcc	gatcgagcgc	gtcaccgggt	gggacacccc	3840
ctaccctcac	gcgcaggaat	gggcttactt	cccagggcct	tcgcgggtag	gtgcggcatt	3900
gaaaaaggtc	atggaggtct	gaatgggcac	gcacgtcatc	aagatgccgg	acattggcga	3960
aggcatcgcg	caggtcgaat	tggtggaatg	gttcgtcaag	gtgggcgaca	tcatcgccga	4020
ggaccaagtg	gtagccgacg	tcatgaccga	caaggccacc	gtggaaatcc	cgtcgccggt	4080
cagcggcaag	gtgctggccc	tgggtggcca	gccaggtgaa	gtgatggcgg	tcggcagtga	4140
gctgatccgc	atcgaagtgg	aaggcagcgg	caaccatgtg	gatgtgccgc	aagccaagcc	4200
ggccgaagtg	cctgcggcac	cggtagccgc	taaacctgaa	ccacagaaaag	acgttaaacc	4260
ggcggcgtag	caggcgtcag	ccagccacga	ggcagcgccc	atcgtgccgc	gccagccggg	4320
cgacaagccg	ctggcctcgc	cggcggtgcg	caaacgcgcc	ctcgatgccg	gcatcgaatt	4380
gcgttatgtg	cacggcagcg	gcccggccgg	gcgcacacct	cacgaagacc	tcgacgcgtt	4440
catgagcaaa	ccgcaaagcg	ctgccgggca	aacccccaat	ggctatgcca	ggcgcaccga	4500
cagcgagcag	gtgccggtga	tcggcctgcg	ccgcaagatc	gcccagcgca	tgcaggacgc	4560
caagcgccgg	gtcgcgcact	tcagctatgt	ggaagaaatc	gacgtcaccg	ccctggaagc	4620
cctgcgccag	cagctcaaca	gcaagcacgg	cgacagccgc	ggcaagctga	cactgctgcc	4680
gttcctggtg	cgcgccctgg	tcgtggcact	gcgtgacttc	ccgcagataa	acgccaccta	4740
cgatgacgaa	gcgcagatca	tcacccgcca	tggcgcggtg	catgtgggca	tcgccaccca	4800
aggtgacaac	ggcctgatgg	taccctgtct	gcgccacgcc	gaagcgggca	gcctgtgggc	4860
caatgccggt	gagatttcac	gcctggccaa	cgctgcgcgc	aacaacaagg	ccagccgcga	4920

20130927_CL5646WOPCT_ST25. txt

```

agagctgtcc ggttcgacca ttaccctgac cagcctcggc gccctgggcg gcatcgtcag 4980
cacgccggtg gtcaacaccc cggaagtggc gatcgtcggg gtcaaccgca tggttgagcg 5040
gcccgtggtg atcgacggcc agatcgtcgt gcgcaagatg atgaacctgt ccagctcggt 5100
cgaccaccgc gtggtcgatg gcatggacgc cgccctgttc atccaggccg tgcgtggcct 5160
gctcgaacaa cccgcctgcc tgttcgtgga gtgagcatgc aacagactat ccagacaacc 5220
ctgttgatca tcggcggcgg ccctggcggc tatgtggcgg ccatccgcgc cgggcaactg 5280
ggcatcccta ccgtgctggt ggaaggccag gcgctgggcg gtacctgcct gaacatcggc 5340
tgcattccgt ccaaggcgct gatccatgtg gccgagcagt tccaccaggc ctgcgcgttt 5400
accgaaccct cgccgctggg catcagcgtg gcttcgccac gcctggacat cggccagagc 5460
gtggcctgga aagacggcat cgtcgatcgc ctgaccactg gtgtcgccgc cctgctgaaa 5520
aagcacgggg tgaagggtgt gcacggctgg gccaaagtgc ttgatggcaa gcaggtcgag 5580
gtggatggcc agcgcattca gtgcgagcac ctgttgctgg ccacgggctc cagcagtgtc 5640
gaactgccga tgctgccgtt ggggtgggccc gtgatttcct cgaccgaggc cctggcaccg 5700
aaagccctgc cgcaacacct ggtggtggtg ggcggtggct acatcggcct ggagctgggt 5760
atcgcctacc gcaagctcgg cgcgcaggtc agcgtggtgg aagcgcgcga gcgcattcctg 5820
ccgacttacg acagcgaact gaccgccccg gtggccgagt cgctgaaaaa gctgggtatc 5880
gccctgcacc ttggccacag cgtcgaaggt tacgaaaatg gctgcctgct ggccaacgat 5940
ggcaagggcg gacaactgcg cctggaagcc gaccgggtgc tgggtggcgt gggccgccgc 6000
ccacgcacca agggcttcaa cctggaatgc ctggacctga agatgaatgg tgccgcgatt 6060
gccatcgacg agcgtgccg gaccagcatg cacaacgtct gggccatcgg cgacgtggcc 6120
ggcgaaccga tgctggcgca cggggccatg gccagggcg agatggtggc cgagatcatc 6180
gccggcaagg cagccgctt cgaaccgct gcgatagccg ccgtgtgctt caccgaccgc 6240
gaagtggctg tggtcggcaa gacgccgga caggccagtc agcaaggcct ggactgcatc 6300
gtcgcgcagt tcccgttcgc cgccaacggc cgggccatga gcctggagtc gaaaagcgg 6360
ttcgtgcgcg tggtcgcgcg gcgtgacaac cacctgatcc tgggctggca agcggttggc 6420
gtggcggttt ccgagctgtc cacggcggtt gccagtcgc tggagatggg tgccctgcctg 6480
gaggatgtgg ccggtacat ccatgcccac ccgaccctgg gtgaagcggg acaggaagcg 6540
gcactgcgtg ccctgggcca cgccctgcat atctgacact gaagcggccg aggccgattt 6600
ggcccgcgc gccgagaggc gctgcgggtc tttttatac ctg 6643

```

<210> 57

<211> 423

<212> PRT

<213> Pseudomonas putida

<400> 57

Met Gly Thr His Val Ile Lys Met Pro Asp Ile Gly Glu Gly Ile Ala

1 5 10 15
 Gln Val Glu Leu Val Glu Trp Phe Val Lys Val Gly Asp Ile Ile Ala
 20 25 30
 Glu Asp Gln Val Val Ala Asp Val Met Thr Asp Lys Ala Thr Val Glu
 35 40 45
 Ile Pro Ser Pro Val Ser Gly Lys Val Leu Ala Leu Gly Gly Gln Pro
 50 55 60
 Gly Glu Val Met Ala Val Gly Ser Glu Leu Ile Arg Ile Glu Val Glu
 65 70 75 80
 Gly Ser Gly Asn His Val Asp Val Pro Gln Ala Lys Pro Ala Glu Val
 85 90 95
 Pro Ala Ala Pro Val Ala Ala Lys Pro Glu Pro Gln Lys Asp Val Lys
 100 105 110
 Pro Ala Ala Tyr Gln Ala Ser Ala Ser His Glu Ala Ala Pro Ile Val
 115 120 125
 Pro Arg Gln Pro Gly Asp Lys Pro Leu Ala Ser Pro Ala Val Arg Lys
 130 135 140
 Arg Ala Leu Asp Ala Gly Ile Glu Leu Arg Tyr Val His Gly Ser Gly
 145 150 155 160
 Pro Ala Gly Arg Ile Leu His Glu Asp Leu Asp Ala Phe Met Ser Lys
 165 170 175
 Pro Gln Ser Ala Ala Gly Gln Thr Pro Asn Gly Tyr Ala Arg Arg Thr
 180 185 190
 Asp Ser Glu Gln Val Pro Val Ile Gly Leu Arg Arg Lys Ile Ala Gln
 195 200 205
 Arg Met Gln Asp Ala Lys Arg Arg Val Ala His Phe Ser Tyr Val Glu
 210 215 220
 Glu Ile Asp Val Thr Ala Leu Glu Ala Leu Arg Gln Gln Leu Asn Ser
 225 230 235 240
 Lys His Gly Asp Ser Arg Gly Lys Leu Thr Leu Leu Pro Phe Leu Val
 245 250 255
 Arg Ala Leu Val Val Ala Leu Arg Asp Phe Pro Gln Ile Asn Ala Thr
 260 265 270
 Tyr Asp Asp Glu Ala Gln Ile Ile Thr Arg His Gly Ala Val His Val
 Page 81

275

280

285

Gly Ile Ala Thr Gln Gly Asp Asn Gly Leu Met Val Pro Val Leu Arg
 290 295 300

His Ala Glu Ala Gly Ser Leu Trp Ala Asn Ala Gly Glu Ile Ser Arg
 305 310 315 320

Leu Ala Asn Ala Ala Arg Asn Asn Lys Ala Ser Arg Glu Glu Leu Ser
 325 330 335

Gly Ser Thr Ile Thr Leu Thr Ser Leu Gly Ala Leu Gly Gly Ile Val
 340 345 350

Ser Thr Pro Val Val Asn Thr Pro Glu Val Ala Ile Val Gly Val Asn
 355 360 365

Arg Met Val Glu Arg Pro Val Val Ile Asp Gly Gln Ile Val Val Arg
 370 375 380

Lys Met Met Asn Leu Ser Ser Ser Phe Asp His Arg Val Val Asp Gly
 385 390 395 400

Met Asp Ala Ala Leu Phe Ile Gln Ala Val Arg Gly Leu Leu Glu Gln
 405 410 415

Pro Ala Cys Leu Phe Val Glu
 420

<210> 58
 <211> 6643
 <212> DNA
 <213> Pseudomonas putida

<400> 58
 gcatgcctgc aggccgccga tgaaatggtg gaaggtatcg gtaggctggc cctgctcatc 60
 gctgaacacg ttacgcccgc tgccggtatc gaccaggctc tggatgaatat gcatggaact 120
 gccaggcgtg cgcgccagcg gtttgcccat gcacaccacg gtcagcccgt gcttgagtgc 180
 cacttccttg agcaggtggt tgaacaggaa ggtctggtcg gccagcagca gcgggtcgcc 240
 atgtagcaag ttgatctcga actggctgac gccatttcg tgcataagg tgcgcgcgg 300
 caggccgagc gcggccatgc actggtacac ctattgaag aacgggcgca ggccgttggt 360
 ggaactgaca ctgaacgccg aatggcccag ctgcgcgcgg ccgtcggtgc ccagcggtgg 420
 ctggaacggc tgctgcgggt cactgttggg ggcaaacacg aagaactcaa gtcggtcgc 480
 cactaccggt gccagacca acgtgcgta gcgggcgatc acggccttca gctggccccg 540
 ggtggacagt gccgagggcc ggcatccag ttattggca tcgcagatgg ccagggcgcg 600
 accgtcatcg ctccaggga agcatgaac ctggctgggt tccgtacca acgccaggtc 660
 gccgtcgtcg cagccgtaga atttcgccg cgggtagccg ccatgatgc attgcagcag 720

20130927_CL5646WOPCT_ST25. txt

cacccacg	gccatctgca	ggcggcgcc	ttcgagaaag	ccttcggcg	tcacacctt	780
gccgcgtggg	acgccgttga	ggtcgggggt	gacgcattcg	atttcacga	tgccctggag	840
ctgagcgatg	ctcatgacgc	ttgtccttgt	tgttgtaggc	tgacaacaac	ataggctggg	900
ggtgtttaaa	atatcaagca	gcctctcgaa	cgcttggggc	ctcttctatt	cgcgcaaggt	960
catgccattg	gccggcaacg	gcaaggctgt	cttgtagcgc	acctgtttca	aggcaaaact	1020
cgagcggata	ttcgccacac	ccggcaaccg	ggtcaggtaa	tcgagaaacc	gctccagcgc	1080
ctggatactc	ggcagcagta	cccgaacag	gtagtccggg	tcgcccgtca	tcaggtagca	1140
ctccatcacc	tcggggcgtt	cggcaatttc	ttcctcgaa	cgggtgcagcg	actgctctac	1200
ctgtttttcc	aggctgacat	ggatgaacac	attcacatcc	agccccaacg	cctcgggcga	1260
caacaaggctc	acctgctggc	ggatcacccc	cagttcttcc	atggcccgc	cccggttgaa	1320
acagggcgtg	ggcgacaggt	tgaccgagcg	tgccagctcg	gcgttggtga	tgcgggcgtt	1380
ttcctgcagg	ctgttgagaa	tgccgatatc	ggtacgatcg	agtttgcgca	tgagacaaaa	1440
tcaccggttt	tttgtgttta	tgcggaatgt	ttatctgccc	cgctcggcaa	aggcaatcaa	1500
cttgagagaa	aaattctcct	gccggaccac	taagatgtag	gggacgctga	cttaccagtc	1560
acaagccggt	actcagcggc	ggccgcttca	gagctcacia	aaacaaatac	ccgagcgagc	1620
gtaaaaagca	tgaacgagta	cgccccctg	cgtttgcatg	tgcccagacc	caccggccgg	1680
ccaggctgcc	agaccgattt	ttcctacctg	cgctgaacg	atgcaggctca	agcccgtaaa	1740
ccccctgtcg	atgtcgacgc	tgccgacacc	gccgacctgt	cctacagcct	ggtccgcgtg	1800
ctcgacgagc	aaggcgacgc	ccaaggcccc	tgggctgaag	acatcgaccc	gcagatcctg	1860
cgccaaggca	tgcgcgccat	gctcaagacg	cggatcttcg	acagccgcat	ggtggttgcc	1920
cagcgccaga	agaagatgtc	cttctacatg	cagagcctgg	gcgaagaagc	catcggcagc	1980
ggccaggcgc	tggcgcttaa	ccgcaccgac	atgtgcttcc	ccacctaccg	tcagcaaagc	2040
atcctgatgg	cccgcgacgt	gtcgctggtg	gagatgatct	gccagttgct	gtccaacgaa	2100
cgcgaccccc	tcaagggccg	ccagctgccg	atcatgtact	cggtagcgga	ggccggcttc	2160
ttcaccatca	gcggcaacct	ggcgaccag	ttcgtgcagg	cggtcggctg	ggccatggcc	2220
tcggcgatca	agggcgatac	caagattgcc	tcggcctgga	tcggcgacgg	cgccactgcc	2280
gaatcggact	tccacaccgc	cctcaccttt	gcccacgttt	accgcgcccc	ggtgatcctc	2340
aacgtggtca	acaaccagtg	ggccatctca	accttcagg	ccatcgccgg	tggcgagtcg	2400
accaccttcg	ccggccgtgg	cgtgggctgc	ggcatcgctt	cgctgcgggt	ggacggcaac	2460
gacttcgtcg	ccgtttacgc	cgcttcgcgc	tgggctgccg	aacgtgccc	ccgtggtttg	2520
ggcccagacc	tgatcgagtg	ggtcacctac	cgtgccggcc	cgactcgac	ctcgacgac	2580
ccgtccaagt	accgccctgc	cgatgactgg	agccacttcc	cgctgggtga	cccgatcgcc	2640
cgctgaagc	agcacctgat	caagatcggc	cactggtccg	aagaagaaca	ccaggccacc	2700
acggccgagt	tcgaagcggc	cgtgattgct	gcgcaaaaag	aagccgagca	gtacggcacc	2760

20130927_CL5646WOPCT_ST25. txt

ctggccaacg	gtcacatccc	gagcgccgcc	tcgatgttcg	aggacgtgta	caaggagatg	2820
cccgaccacc	tgcgccgcc	acgccaggaa	ctgggggttt	gagatgaacg	accacaacaa	2880
cagcatcaac	ccggaaccg	ccatggccac	cactacatg	accatgatcc	aggccctgcg	2940
ctcggccatg	gatgtcatgc	ttgagcgca	cgacaatgtg	gtggtgtacg	gccaggacgt	3000
cggctacttc	ggcggcgtgt	tccgctgcac	cgaaggcctg	cagaccaagt	acggcaagtc	3060
ccgcgtgttc	gacgcgccc	tctctgaaag	cggcatcgtc	ggcaccgccg	tgggcatggg	3120
tgcctacggc	ctgcgcccgg	tggtggaat	ccagttcgct	gactacttct	acccggcctc	3180
cgaccagatc	gtttctgaaa	tggcccgct	gcgctaccgt	tcggccggcg	agttcatcgc	3240
cccgtgacc	ctgcgtatgc	cctgcggtgg	cggatatctat	ggcggccaga	cacacagcca	3300
gagcccggaa	gcgatgttca	ctcaggtgtg	cggcctgcgc	accgtaatgc	catccaaccc	3360
gtacgacgcc	aaaggcctgc	tgattgcctc	gatcgaatgc	gacgaccgg	tgatcttct	3420
ggagcccaag	cgcctgtaca	acggcccgtt	cgacggccac	catgaccgcc	cggttacgcc	3480
gtggtcgaaa	cacccgcaca	gcgccgtgcc	cgatggctac	tacaccgtgc	cactggacaa	3540
ggccgccatc	acccgccccg	gcaatgacgt	gagcgtgctc	acctatggca	ccaccgtgta	3600
cgtggcccag	gtggccgccg	aagaaagtgg	cgtggatgcc	gaagtgatcg	acctgcgcag	3660
cctgtggccg	ctagacctgg	acaccatcgt	cgagtcggtg	aaaaagaccg	gccgttgct	3720
ggtagtacac	gaggccaccc	gtacttgttg	ctttggcgca	gaactggtgt	cgctggtgca	3780
ggagcactgc	ttccaccacc	tggaggcgcc	gatcgagcgc	gtcaccggtt	gggacacccc	3840
ctaccctcac	gcgcaggaat	gggcttactt	cccagggcct	tcgcgggtag	gtgcggcatt	3900
gaaaaaggtc	atggaggctt	gaatgggcac	gcacgtcatc	aagatgccgg	acattggcga	3960
aggcatcgcg	caggtcgaat	tggtggaatg	gttcgtcaag	gtgggcgaca	tcatcgccga	4020
ggaccaagtg	gtagccgacg	tcatgaccga	caaggccacc	gtggaaatcc	cgtcgccggt	4080
cagcggcaag	gtgctggccc	tgggtggcca	gccaggtgaa	gtgatggcgg	tcggcagtga	4140
gctgatccgc	atcgaagtgg	aaggcagcgg	caaccatgtg	gatgtgccgc	aagccaagcc	4200
ggccgaagtg	cctgcggcac	cggtagccgc	taaacctgaa	ccacagaaaag	acgttaaacc	4260
ggcggcgtag	caggcgtcag	ccagccacga	ggcagcgccc	atcgtgccgc	gccagccggg	4320
cgacaagccg	ctggcctcgc	cggcgggtcg	caaacgcgcc	ctcgatgccg	gcacgaatt	4380
gcgttatgtg	cacggcagcg	gcccggccgg	gcgcacctg	cacgaagacc	tcgacgcgtt	4440
catgagcaaa	ccgcaaagcg	ctgccgggca	aacccccaat	ggctatgcca	ggcgcaccga	4500
cagcgagcag	gtgccggtga	tcggcctgcg	ccgcaagatc	gccagcgca	tgcaggacgc	4560
caagcgccgg	gtcgcgcact	tcagctatgt	ggaagaaatc	gacgtcaccg	ccctggaagc	4620
cctgcgccag	cagctcaaca	gcaagcacgg	cgacagccgc	ggcaagctga	cactgctgcc	4680
gttcctggtg	cgcgccctgg	tcgtggcact	gcgtgacttc	ccgcagataa	acgccaccta	4740
cgatgacgaa	gcgcagatca	tcacccgcca	tggcgcggtg	catgtgggca	tcgccaccca	4800

20130927_CL5646WOPCT_ST25. txt

```

aggtgacaac ggcctgatgg taccctgtct gcgccacgcc gaagcgggca gcctgtgggc 4860
caatgccggt gagatttcac gcctggccaa cgctgcgcgc aacaacaagg ccagccgcga 4920
agagctgtcc ggttcgacca ttaccctgac cagcctcggc gccctgggcg gcatcgtcag 4980
cacgccggtg gtcaacaccc cggaagtggc gatcgtcggg gtcaaccgca tggttgagcg 5040
gcccgtgggtg atcgacggcc agatcgtcgt gcgcaagatg atgaacctgt ccagctcgtt 5100
cgaccaccgc gtggtcgatg gcatggacgc cgccctgttc atccaggccg tgcgtggcct 5160
gctcgaacaa cccgcctgcc tgttcgtgga gtgagcatgc aacagactat ccagacaacc 5220
ctgttgatca tcggcggcgg ccctggcggc tatgtggcgg ccatccgcgc cgggcaactg 5280
ggcatcccta ccgtgctggt ggaaggccag gcgctgggcg gtacctgcct gaacatcggc 5340
tgcatccgt ccaaggcgct gatccatgtg gccgagcagt tccaccaggc ctgcgcgttt 5400
accgaaccct cgccgctggg catcagcgtg gcttcgccac gcctggacat cggccagagc 5460
gtggcctgga aagacggcat cgtcgatcgc ctgaccactg gtgtcgccgc cctgctgaaa 5520
aagcacgggg tgaagggtgt gcacggctgg gccaaagggtc ttgatggcaa gcaggtcgag 5580
gtggatggcc agcgcaccca gtgcgagcac ctgttgctgg ccacgggctc cagcagtgtc 5640
gaactgccga tgctgccgtt ggggtgggccc gtgatttcct cgaccgaggc cctggcaccg 5700
aaagccctgc cgcaacacct ggtggtggtg ggcggtggct acatcggcct ggagctgggt 5760
atgcctacc gcaagctcgg cgcgcaggtc agcgtggtgg aagcgcgcga gcgcatcctg 5820
ccgacttacg acagcgaact gaccgccccg gtggccgagt cgctgaaaaa gctgggtatc 5880
gccctgcacc ttggccacag cgtcgaaggt tacgaaaatg gctgcctgct ggccaacgat 5940
ggcaagggcg gacaactgcg cctggaagcc gaccgggtgc tgggtggcgt gggccgccgc 6000
ccacgcacca agggcttcaa cctggaatgc ctggacctga agatgaatgg tgccgcgatt 6060
gccatcgacg agcgtgccca gaccagcatg cacaacgtct gggccatcgg cgacgtggcc 6120
ggcgaaccga tgctggcgca ccgggcatg gccagggcg agatggtggc cgagatcatc 6180
gccggcaagg cacgccgtt cgaacccgt gcgatagccg ccgtgtgctt caccgaccgc 6240
gaagtggctg tggtcggcaa gacgccgga caggccagtc agcaaggcct ggactgcatc 6300
gtcgcgcagt tccgttcgc cgccaacggc cgggccatga gcctggagtc gaaaagcggg 6360
ttcgtgcgcg tggtcgcgcg gcgtgacaac cacctgatcc tgggctggca agcggttggc 6420
gtggcggttt ccgagctgtc cacggcgttt gccagtcgc tggagatggg tgcctgcctg 6480
gaggatgtgg ccggtacat ccatgccac ccgaccctgg gtgaagcggg acaggaagcg 6540
gcactgcgtg ccctgggcca cgccctgcat atctgacact gaagcggccg aggccgattt 6600
ggcccggcgc gccgagaggc gctgcgggtc tttttatata ctg 6643

```

```

<210> 59
<211> 459
<212> PRT
<213> Pseudomonas putida

```

<400> 59

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

260

265

270

Gly Phe Asn Leu Glu Cys Leu Asp Leu Lys Met Asn Gly Ala Ala Ile
 275 280 285

Ala Ile Asp Glu Arg Cys Gln Thr Ser Met His Asn Val Trp Ala Ile
 290 300

Gly Asp Val Ala Gly Glu Pro Met Leu Ala His Arg Ala Met Ala Gln
 305 310 315 320

Gly Glu Met Val Ala Glu Ile Ile Ala Gly Lys Ala Arg Arg Phe Glu
 325 330 335

Pro Ala Ala Ile Ala Ala Val Cys Phe Thr Asp Pro Glu Val Val Val
 340 345 350

Val Gly Lys Thr Pro Glu Gln Ala Ser Gln Gln Gly Leu Asp Cys Ile
 355 360 365

Val Ala Gln Phe Pro Phe Ala Ala Asn Gly Arg Ala Met Ser Leu Glu
 370 375 380

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

Ile Leu Gly Trp Gln Ala Val Gly Val Ala Val Ser Glu Leu Ser Thr
 405 410 415

Ala Phe Ala Gln Ser Leu Glu Met Gly Ala Cys Leu Glu Asp Val Ala
 420 425 430

Gly Thr Ile His Ala His Pro Thr Leu Gly Glu Ala Val Gln Glu Ala
 435 440 445

Ala Leu Arg Ala Leu Gly His Ala Leu His Ile
 450 455

<210> 60

<211> 6643

<212> DNA

<213> Pseudomonas putida

<400> 60

gcatgcctgc aggccgccga tgaaatggtg gaaggtatcg gtaggctggc cctgctcatc 60

gctgaacacg ttacgcccgc tgccggtatc gaccaggctc tggatgaatat gcatggaact 120

gccaggcgtg cgcgccagcg gtttggccat gcacaccacg gtcagcccgt gcttgagtgc 180

cacttccttg agcagggtgtt tgaacaggaa ggtctggtcg gccagcagca gcgggtcgcc 240

atgtagcaag ttgatctcga actggctgac gccatttcg tgcataagg tgcgcgcgg 300

caggccgagc gcggccatgc actggtacac ctattgaag aacgggcgca ggccgttggt 360

20130927_CL5646WOPCT_ST25. txt

ggaactgaca	ctgaacgccg	aatggcccag	ctcgcggcgg	ccgtcgggtgc	ccagcgggtgg	420
ctggaacggc	tgctgcgggt	cactgtttgg	ggcaaacacg	aagaactcaa	gctcgggtcgc	480
cactaccggt	gccagaccca	acgctgcgta	gcgggcgatc	acggccttca	gctggccccg	540
ggtggacagt	gccgagggcc	ggccatccag	ttcattggca	tcgcagatgg	ccagggcgcg	600
accgtcatcg	ctccagggga	agcgatgaac	ctggctgggt	tccgctacca	acgccaggtc	660
gccgtcgtcg	cagccgtaga	atttcgccgg	cgggtagccg	cccatgatgc	attgcagcag	720
cacccacagg	gccatctgca	ggcggcgggc	ttcagaaaag	ccttcggcgg	tcatacctt	780
gccgcgtggg	acgccgttga	ggtcgggggt	gacgcattcg	atttcatcga	tgccctggag	840
ctgagcgatg	ctcatgacgc	ttgtccttgt	tgttgtaggc	tgacaacaac	ataggctggg	900
ggtgtttaaa	atatcaagca	gcctctcgaa	cgcctggggc	ctcttctatt	cgcgcaagggt	960
catgccattg	gccggcaacg	gcaaggctgt	cttgtagcgc	acctgtttca	aggcaaaact	1020
cgagcggata	ttcgccacac	ccggcaaccg	ggtcaggtaa	tcgagaaacc	gctccagcgc	1080
ctggatactc	ggcagcagta	cccgaacag	gtagtccggg	tcgcccgtca	tcaggtagca	1140
ctccatcacc	tcggggccgtt	cggcaatttc	ttcctcgaag	cgggtgcagcg	actgctctac	1200
ctgtttttcc	aggctgacat	ggatgaacac	attcacatcc	agccccaacg	cctcgggcga	1260
caacaaggtc	acctgctggc	ggatcacccc	cagttcttcc	atggcccgca	cccggttgaa	1320
acagggcggtg	ggcgacaggt	tgaccgagcg	tgccagctcg	gcgttggtga	tgcgggcggtt	1380
ttcctgcagg	ctgttgagaa	tgccgatatc	ggtacgatcg	agtttgcgca	tgagacaaaa	1440
tcaccggttt	tttgtgttta	tgcggaatgt	ttatctgccc	cgctcggcaa	aggcaatcaa	1500
cttgagagaa	aaattctcct	gccggaccac	taagatgtag	gggacgctga	cttaccagtc	1560
acaagccggt	actcagcggc	ggccgcttca	gagctcacia	aaacaaatac	ccgagcgagc	1620
gtaaaaagca	tgaacgagta	cgccccctg	cgtttgcatg	tgcccagagcc	caccggccgg	1680
ccagggtgcc	agaccgattt	ttcctacctg	cgcctgaacg	atgcagggtca	agcccgtaaa	1740
ccccctgtcg	atgtcgacgc	tgccgacacc	gccgacctgt	cctacagcct	ggtccgcgtg	1800
ctcgacgagc	aaggcgacgc	ccaaggcccc	tgggctgaag	acatcgaccc	gcagatcctg	1860
cgccaaggca	tgcgcgccat	gctcaagacg	cggatcttcg	acagccgcat	ggtggttgcc	1920
cagcgccaga	agaagatgtc	cttctacatg	cagagcctgg	gcgaagaagc	catcggcagc	1980
ggccaggcgc	tggcgcttaa	ccgcaccgac	atgtgcttcc	ccacctaccg	tcagcaaagc	2040
atcctgatgg	cccgcgacgt	gtcgctggtg	gagatgatct	gccagttgct	gtccaacgaa	2100
cgcgaccccc	tcaagggccg	ccagctgccg	atcatgtact	cggtagcgga	ggccggcttc	2160
ttcaccatca	gcggcaacct	ggcgacccag	ttcgtgcagg	cggtcggctg	ggccatggcc	2220
tcggcgatca	agggcgatac	caagattgcc	tcggcctgga	tcggcgacgg	cgccactgcc	2280
gaatcggact	tccacaccgc	cctcaccttt	gcccacgttt	accgcgcccc	ggtgatcctc	2340
aacgtggtca	acaaccagtg	ggccatctca	accttcagg	ccatcgccgg	tggcgagtcg	2400

20130927_CL5646WOPCT_ST25. txt

accaccttcg	ccggccgtgg	cgtgggctgc	ggcatcgctt	cgctgcgggt	ggacggcaac	2460
gacttcgtcg	ccgtttacgc	cgcttcgcgc	tgggctgccg	aacgtgcccg	ccgtggtttg	2520
ggcccagacc	tgatcgagtg	ggtcacctac	cgtgccggcc	cgcactcgac	ctcggacgac	2580
ccgtccaagt	accgccctgc	cgatgactgg	agccacttcc	cgctgggtga	cccgatcgcc	2640
cgcctgaagc	agcacctgat	caagatcggc	cactgggtccg	aagaagaaca	ccaggccacc	2700
acggccgagt	tcgaagcggc	cgtgattgct	gcgcaaaaag	aagccgagca	gtacggcacc	2760
ctggccaacg	gtcacatccc	gagcgccgcc	tcgatgttcg	aggacgtgta	caaggagatg	2820
cccgaccacc	tgcgccgcca	acgccaggaa	ctgggggttt	gagatgaacg	accacaacaa	2880
cagcatcaac	ccggaaaccg	ccatggccac	cactaccatg	accatgatcc	aggccctgcg	2940
ctcggccatg	gatgtcatgc	ttgagcgcg	cgacaatgtg	gtggtgtacg	gccaggacgt	3000
cggctacttc	ggcggcgtgt	tccgctgcac	cgaaggcctg	cagaccaagt	acggcaagtc	3060
ccgcgtgttc	gacgcgcccc	tctctgaaag	cggcatcgtc	ggcaccgccg	tgggcatggg	3120
tgcctacggc	ctgcgcccgg	tggtggaat	ccagttcgct	gactacttct	acccggcctc	3180
cgaccagatc	gtttctgaaa	tggcccgctt	gcgctaccgt	tcggccggcg	agttcatcgc	3240
cccgtgacc	ctgcgtatgc	cctgcgggtg	cggatatctat	ggcggccaga	cacacagcca	3300
gagcccggaa	gcgatgttca	ctcagggtgtg	cggcctgcgc	accgtaatgc	catccaaccc	3360
gtacgacgcc	aaaggcctgc	tgattgcctc	gatcgaatgc	gacgaccggg	tgatcttctt	3420
ggagcccaag	cgctgtaca	acggcccgtt	cgacggccac	catgaccgcc	cggttacgcc	3480
gtggtcgaaa	cacccgcaca	gcgccgtgcc	cgatggctac	tacaccgtgc	cactggacaa	3540
ggccgccatc	acccgccccg	gcaatgacgt	gagcgtgctc	acctatggca	ccaccgtgta	3600
cgtggcccag	gtggccggcg	aagaaagtgg	cgtggatgcc	gaagtgatcg	acctgcgcag	3660
cctgtggccg	ctagacctgg	acaccatcgt	cgagtcgggtg	aaaaagaccg	gccgttgctg	3720
ggtagtacac	gaggccaccc	gtacttgtgg	ctttggcgca	gaactggtgt	cgctggtgca	3780
ggagcactgc	ttccaccacc	tggaggcgcc	gatcgagcgc	gtcaccgggtt	gggacacccc	3840
ctaccctcac	gcgcaggaat	gggcttactt	cccagggcct	tcgcgggtag	gtgcggcatt	3900
gaaaaaggtc	atggaggctt	gaatgggcac	gcacgtcatc	aagatgccgg	acattggcga	3960
aggcatcgcg	caggtcgaat	tggtggaatg	gttcgtcaag	gtgggcgaca	tcacgcgccg	4020
ggaccaagtg	gtagccgacg	tcatgaccga	caaggccacc	gtggaaatcc	cgtcgccgggt	4080
cagcggcaag	gtgctggccc	tgggtggcca	gccaggtgaa	gtgatggcgg	tcggcagtga	4140
gctgatccgc	atcgaagtgg	aaggcagcgg	caaccatgtg	gatgtgccgc	aagccaagcc	4200
ggccgaagtg	cctgcggcac	cggtagccgc	taaacctgaa	ccacagaaaag	acgttaaacc	4260
ggcggcgtag	caggcgtagc	ccagccacga	ggcagcgccc	atcgtgccgc	gccagccggg	4320
cgacaagccg	ctggcctcgc	cggcggtgcg	caaacgcgcc	ctcgatgccg	gcacgaattt	4380
gcgttatgtg	cacggcagcg	gcccggccgg	gcgcatcctg	cacgaagacc	tcgacgcgtt	4440

20130927_CL5646WOPCT_ST25. txt

catgagcaaa	ccgcaaagcg	ctgccgggca	aacccccaat	ggctatgcca	ggcgcaccga	4500
cagcgagcag	gtgccggtga	tcggcctgcg	ccgcaagatc	gcccagcgca	tgcaggacgc	4560
caagcgccgg	gtcgcgcact	tcagctatgt	ggaagaaatc	gacgtcaccg	ccctggaagc	4620
cctgcgccag	cagctcaaca	gcaagcacgg	cgacagccgc	ggcaagctga	cactgctgcc	4680
gttcctggtg	cgcgccctgg	tcgtggcact	gcgtgacttc	ccgcagataa	acgccaccta	4740
cgatgacgaa	gcgcagatca	tcacccgcca	tggcgcggtg	catgtgggca	tcgccaccca	4800
aggtgacaac	ggcctgatgg	tacccgtgct	gcgccacgcc	gaagcgggca	gcctgtgggc	4860
caatgccggt	gagatttcac	gcctggccaa	cgctgcgcgc	aacaacaagg	ccagccgcga	4920
agagctgtcc	ggttcgacca	ttaccctgac	cagcctcggc	gccctgggcg	gcacgtcag	4980
cacgccggtg	gtcaacaccc	cgaagtggc	gatcgtcgg	gtcaaccgca	tggttgagcg	5040
gcccggtgtg	atcgacggcc	agatcgtcgt	gcgcaagatg	atgaacctgt	ccagctcgtt	5100
cgaccaccgc	gtggtcgatg	gcatggacgc	cgccctgttc	atccaggccg	tgcgtggcct	5160
gctcgaacaa	cccgcctgcc	tgttcgtgga	gtgagcatgc	aacagactat	ccagacaacc	5220
ctgttgatca	tcggcggcgg	ccctggcggc	tatgtggcgg	ccatccgcgc	cgggcaactg	5280
ggcatcccta	ccgtgctggt	ggaaggccag	gcgctgggcg	gtacctgcct	gaacatcggc	5340
tgcatccgt	ccaaggcgct	gatccatgtg	gccgagcagt	tccaccaggc	ctcgcgcttt	5400
accgaaccct	cgccgctggg	catcagcgtg	gcttcgccac	gcctggacat	cggccagagc	5460
gtggcctgga	aagacggcat	cgtcgatcgc	ctgaccactg	gtgtcgccgc	cctgctgaaa	5520
aagcacgggg	tgaagggtgt	gcacggctgg	gccaagggtg	ttgatggcaa	gcaggtcgag	5580
gtggatggcc	agcgcatcca	gtgcgagcac	ctgttgctgg	ccacgggctc	cagcagtgtc	5640
gaactgccga	tgctgccgtt	gggtgggccc	gtgatttcct	cgaccgaggc	cctggcaccg	5700
aaagccctgc	cgcaacacct	ggtggtggtg	ggcgtggtg	acatcggcct	ggagctgggt	5760
atgcctacc	gcaagctcgg	cgcgcaggtc	agcgtggtgg	aagcgcgcga	gcgcacacct	5820
ccgacttacg	acagcgaact	gaccgccccg	gtggccgagt	cgctgaaaaa	gctgggtatc	5880
gccctgcacc	ttggccacag	cgtcgaaggt	tacgaaaatg	gctgcctgct	ggccaacgat	5940
ggcaagggcg	gacaactgcg	cctggaagcc	gaccgggtgc	tggtagccgt	gggccgccgc	6000
ccacgcacca	agggttcaa	cctggaatgc	ctggacctga	agatgaatgg	tgccgcgatt	6060
gccatcgacg	agcgtgccca	gaccagcatg	cacaacgtct	gggccatcgg	cgacgtggcc	6120
ggcgaaccga	tgctggcgca	ccgggccatg	gccagggcg	agatggtggc	cgagatcatc	6180
gccggcaagg	cacgccgctt	cgaacccgct	gcgatagccg	ccgtgtgctt	caccgaccgc	6240
gaagtggctg	tggtcggcaa	gacgccggaa	caggccagtc	agcaaggcct	ggactgcac	6300
gtcgcgcagt	tcccgttcgc	cgccaacggc	cgggccatga	gcctggagtc	gaaaagcggt	6360
ttcgtgcgcg	tggtcgcgcg	gcgtgacaac	cacctgatcc	tgggctggca	agcggttggc	6420
gtggcgggtt	ccgagctgtc	cacggcggtt	gccagtcgc	tggagatggg	tgctgcctg	6480

gaggatgtgg ccggtacat ccatgcccac ccgaccctgg gtgaagcggg acaggaagcg 6540
gcactgcgtg ccctgggcca cgccctgcat atctgacact gaagcggccg aggccgattt 6600
ggcccgccgc gccgagaggc gctgcgggtc tttttatac ctg 6643

<210> 61
<211> 468
<212> PRT
<213> Clostridium beijerinckii

<400> 61

Met Asn Lys Asp Thr Leu Ile Pro Thr Thr Lys Asp Leu Lys Leu Lys
1 5 10 15

Thr Asn Val Glu Asn Ile Asn Leu Lys Asn Tyr Lys Asp Asn Ser Ser
20 25 30

Cys Phe Gly Val Phe Glu Asn Val Glu Asn Ala Ile Asn Ser Ala Val
35 40 45

His Ala Gln Lys Ile Leu Ser Leu His Tyr Thr Lys Glu Gln Arg Glu
50 55 60

Lys Ile Ile Thr Glu Ile Arg Lys Ala Ala Leu Glu Asn Lys Glu Val
65 70 75 80

Leu Ala Thr Met Ile Leu Glu Glu Thr His Met Gly Arg Tyr Glu Asp
85 90 95

Lys Ile Leu Lys His Glu Leu Val Ala Lys Tyr Thr Pro Gly Thr Glu
100 105 110

Asp Leu Thr Thr Thr Ala Trp Ser Gly Asp Asn Gly Leu Thr Val Val
115 120 125

Glu Met Ser Pro Tyr Gly Val Ile Gly Ala Ile Thr Pro Ser Thr Asn
130 135 140

Pro Thr Glu Thr Val Ile Cys Asn Ser Ile Gly Met Ile Ala Ala Gly
145 150 155 160

Asn Ala Val Val Phe Asn Gly His Pro Gly Ala Lys Lys Cys Val Ala
165 170 175

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

Glu Asn Leu Val Thr Thr Ile Lys Asn Pro Thr Met Glu Ser Leu Asp
195 200 205

Ala Ile Ile Lys His Pro Leu Ile Lys Leu Leu Cys Gly Thr Gly Gly
Page 91

210

215

220

Pro Gly Met Val Lys Thr Leu Leu Asn Ser Gly Lys Lys Ala Ile Gly
 225 230 235 240

Ala Gly Ala Gly Asn Pro Pro Val Ile Val Asp Asp Thr Ala Asp Ile
 245 250 255

Glu Lys Ala Gly Lys Ser Ile Ile Glu Gly Cys Ser Phe Asp Asn Asn
 260 265 270

Leu Pro Cys Ile Ala Glu Lys Glu Val Phe Val Phe Glu Asn Val Ala
 275 280 285

Asp Asp Leu Ile Ser Asn Met Leu Lys Asn Asn Ala Val Ile Ile Asn
 290 295 300

Glu Asp Gln Val Ser Lys Leu Ile Asp Leu Val Leu Gln Lys Asn Asn
 305 310 315 320

Glu Thr Gln Glu Tyr Phe Ile Asn Lys Lys Trp Val Gly Lys Asp Ala
 325 330 335

Lys Leu Phe Ser Asp Glu Ile Asp Val Glu Ser Pro Ser Asn Ile Lys
 340 345 350

Cys Ile Val Cys Glu Val Asn Ala Asn His Pro Phe Val Met Thr Glu
 355 360 365

Leu Met Met Pro Ile Leu Pro Ile Val Arg Val Lys Asp Ile Asp Glu
 370 375 380

Ala Val Lys Tyr Thr Lys Ile Ala Glu Gln Asn Arg Lys His Ser Ala
 385 390 395 400

Tyr Ile Tyr Ser Lys Asn Ile Asp Asn Leu Asn Arg Phe Glu Arg Glu
 405 410 415

Ile Asp Thr Thr Ile Phe Val Lys Asn Ala Lys Ser Phe Ala Gly Val
 420 425 430

Gly Tyr Glu Ala Glu Gly Phe Thr Thr Phe Thr Ile Ala Gly Ser Thr
 435 440 445

Gly Glu Gly Ile Thr Ser Ala Arg Asn Phe Thr Arg Gln Arg Arg Cys
 450 455 460

Val Leu Ala Gly
 465

<210> 62

<211> 6558

<212> DNA

<213> Cl ostridi um bei j eri ncki i

<400> 62

```

aagcttaaaa tatccatagg ctattgttaa taagactata gcgcttaata ctctaagcgc      60
accatctaaa aaattataca taggaattgg ataaactcca atcttctcca taatactttt      120
catagggtgaa attgatatta taattaaggc tgcataaagc aaactcatat ctccaataaa      180
tgtcactatc ataggttaatt tccttttcat gttcacatac tcccccgat tctataataa      240
tttacaataa acttccatca caaataaatt ataacatata ttgtaaatat agttttatta      300
ttcgcatatt tatagataaa caatatataa cttagactta tgcaataacc taccgaaagt      360
aaaaacattg ttatttcaca ggactatgaa aatttcgctg aaagtactaa attgcagggt      420
gcaccactaa tgcttgctcc caatttcatt gtgacaagca ttagttgaac aacctacaat      480
taagagcctt taacagctca tttccaatgc ctgctacata aaaatgtttc tactttctaa      540
tatggtatctt acttcaaagt gtaaatacaa ttcttaagtt gtttatctat atagggttcc      600
atattgataa caatacatat ctactgtttc aatttcatga ataccagcga tattgaaatt      660
ttgtaagttt gaatatcatt gtgaaaccct atatatcaaa tacaatctca aaattataca      720
aaaagatccc aacttcacaa tatgaatttg agatcttcta tcttaacttc ttcaatatct      780
ttaagattat ttatactgcg tgcaaatctt ctataagtat caattatcag ctatgtacat      840
tgatagtggg cttgtaatct gaacagctta tatattttaca cttttaagtt ctcttccact      900
aacccttgct ttgaaacttc tcataaacia atcgaaagta cctttaatta gctgtgcatt      960
ctcaaactca gaatttaact taagtacttc aattgccgct tcgatagtat agagctctcc     1020
ttctgaagca ctttttctta aagtatatct tgatttatta attggattta atgaaattct     1080
tggaagcttc ttttaagtagt cgctctttct tagtatcttc tctgcttctt tccatgtgcc     1140
atctaagata ataaatgctg gaattttttc tgaatcttta catttagact ttctttctag     1200
aggttcatca tcatccatag gaaataatac acgtatttca taatcatcgc tattaatata     1260
ttcaattaat ttttcaggag tctttactct ctcccaaaga attaactcag ttgattctgg     1320
attcaccaat ttcaataatc tagcggtatt tgaaggccta ctaaattctc tttctgttga     1380
taatatcaat atctttgctt ttgtctctat tttaggcaca atatcgcaga tacaatttat     1440
tattggcaac ccacatttat tgcagctctc atataactta gtaatttgct taactttaaa     1500
ttcagactcc attttacctc cattattagt tggttagtgt gtcatatctt cttgctatta     1560
ctaactgatt ataacatag tattcaatat atcactccta gttttcaaag cactggcaat     1620
acgaattaca aattaatttc tggatttatg tcagtatttc attaataaaa ggtcggactt     1680
ttaagatact tgtttttagct attgatcata ttatttaaag actatgcatt taatgtataa     1740
ttataatgaa tattatcaat aatatttatt ttatattaca atcttacagt ctttattcta     1800
aatttcactc aaataccaaa cgagctttat tcataaacia tatataacia taattccaaa     1860
ataatacgat attttatctg taacagccat ataaaaaaaa tatcatatag tcttgtcatt     1920

```

20130927_CL5646WOPCT_ST25. txt

tgataacggtt ttgtcttcct tatatttact ttttcggttt aatagggttga ttctgtaaat	1980
tttagtgata acatatatatt gatgacatta aaaatttaaat atttcatata aattttttaat	2040
gtctattaat ttttaaatca caaggaggaa tagttcatga ataaagacac actaatacct	2100
acaactaaag attttaaatt aaaaacaaat gttgaaaaca ttaattttaaa gaactacaag	2160
gataattctt catgtttcgg agtatttcgaa aatgttgaaa atgctataaaa cagcgctgta	2220
cacgcgcaaa agatattatc ccttcattat acaaaaagaac aaagagaaaa aatcataact	2280
gagataagaa aggccgcatt agaaaataaa gaggttttag ctaccatgat tctggaagaa	2340
acacatatgg gaaggtatga agataaaata ttaaagcatg aattagtagc taaatatact	2400
cctggtacag aagatttaac tactactgct tggtcagggtg ataatggtct tacagttgta	2460
gaaatgtctc catatggcgt tatagggtgca ataactcctt ctacgaatcc aactgaaact	2520
gtaatatgta atagcatcgg catgatagct gctggaaatg ctgtagtatt taacggacac	2580
ccaggcgcta aaaaatgtgt tgcttttgct attgaaatga taaataaagc aattattttca	2640
tgtggcggtc ctgagaattt agtaacaact ataaaaaatc caactatgga atccctagat	2700
gcaattatta agcatccttt aataaaaactt ctttgcggaa ctggagggtcc aggaatggta	2760
aaaaccctct taaattctgg caagaaagct atagggtgctg gtgctggaaa tccaccagtt	2820
attgtagatg ataccgctga tatagaaaag gctggtaaga gtatcattga aggctgttct	2880
tttgataata atttaccttg tattgcagaa aaagaagtat ttgtttttga gaatgttgca	2940
gatgatttaa tatctaacat gctaaaaaat aatgctgtaa ttataaatga agatcaagta	3000
tcaaaattaa tagatttagt attacaaaaa aataatgaaa ctcaagaata ctttataaac	3060
aaaaaatggg taggaaaaga tgcaaaatta ttctcagatg aaatagatgt tgagtctcct	3120
tcaaatatta aatgcatagt ctgcgaagta aatgcaaadc atccatttgt catgacagaa	3180
ctcatgatgc caatattacc aattgtaaga gttaaagata tagatgaagc tgttaaatat	3240
acaaagatag cagaacaaaa tagaaaacat agtgcctata ttatttctaa aaatatagac	3300
aacctaaata gatttgaaag agaaattgat actactatit ttgtaaagaa tgctaaatct	3360
tttgctggtg ttggttatga agctgaagga tttaacaact tcactattgc tggatctact	3420
ggtgaaggca taacctctgc aagaaatttt acaagacaaa gaagatgtgt acttgccggc	3480
taacttcttg ctaaatttat acatttattc acataacttt aatatgcaat gttcccacaa	3540
aatattaaaa actatttaga agggagatat taaatgaata aattagtaaa attaacagat	3600
ttaaagcgca ttttcaaaga tggtatgaca attatggttg ggggtttttt agattgtgga	3660
actcctgaaa atattataga tatgctagtt gatttaaata taaaaaatct gactattata	3720
agcaatgata cagcttttcc taataaagga ataggaaaac ttattgtaaa tgggtcaagtt	3780
tctaaagtaa ttgcttcaca tattggaact aatcctgaaa ctgggaaaaa aatgagctct	3840
ggtgaactta aagttgagct ttctccacaa ggaacactga tcgaaagaat tcgtgcagct	3900
ggatctggac tcggagggtgt attaactcca accggacttg ggactatcgt tgaagaaggt	3960

20130927_CL5646WOPCT_ST25. txt

aagaaaaaag	ttactatcgg	tggcaaagaa	tatctattag	aacttccttt	atccgctgat	4020
gtttcattaa	taaaaggtag	cattgtagat	gaatttggaa	ataccttcta	tagagctgct	4080
actaaaaatt	tcaatccata	tatggcaatg	gctgcaaaaa	cagttatagt	tgaagcagaa	4140
aatttagtta	aatgtgaaga	tttaaaaaga	gatgccataa	tgactcctgg	cgtattagta	4200
gattatatcg	ttaaggaggc	ggcttaattg	attgtagata	aagtttttagc	aaaagagata	4260
attgccaaaa	gagttgcaa	agaactaaaa	aaaggccaac	tcgtaaacct	tggaatagga	4320
cttccaactt	tagtagctaa	ttatgtgcca	aaagaaatga	acattacttt	cgaatcagaa	4380
aatggcatgg	ttggcatggc	acaaatggcc	tcatcagggtg	aaaatgaccc	agatataata	4440
aatgctggtg	gggaatatgt	aacattatta	cctcaagggtg	catttttttga	tagttcaacg	4500
tcttttgcac	taataagagg	aggacatggt	gatgttgctg	ttcttggtgc	tctagaagtt	4560
gatgaagaag	gtaatttagc	taactggatt	gttccaaata	aaattgtccc	aggtatggga	4620
ggcgccatgg	atttggaat	aggcgcaaaa	aaaataatag	tggcaatgca	acatacagga	4680
aaaggtaaac	ctaaaatcgt	aaaaaaatgt	actctccac	ttactgctaa	ggctcaggta	4740
gatttaattg	ttacagaact	ttgtgtaatt	gatgtaacaa	atgatggttt	acttttcaga	4800
gaaattcata	aagatacaac	tattgatgaa	ataaaatttt	taacagatgc	agatttaatt	4860
attcccgaca	acttaaaaat	tatggatatc	taaatcattc	tattttaaat	atataacttt	4920
aaaaatctta	tgtattaaaa	actaagaaaa	gaggttgatt	attttatggt	agaaagtga	4980
gtatctaaac	aaattacaac	tccacttgct	gctccagcgt	ttcctagagg	accatataga	5040
tttcacaata	gagaatatct	aaacattatt	tatcgaactg	atttagatgc	tcttcgaaaa	5100
atagtaccag	agccacttga	attagatgga	gcatatgtta	ggtttgagat	gatggctatg	5160
cctgatacaa	ccggactagg	ctcatatact	gagtgtggtc	aagccattcc	agtaaaatat	5220
aatgaggtta	aagggtgacta	cttgcatatg	atgtacctag	ataatgaacc	tgctattgct	5280
gttgaagag	aaagcagtgc	ttatcccaaa	aagttcggct	atccaaagct	atttgttgat	5340
tcagacgccc	tagttggcgc	ccttaagtat	ggtgcattac	cggtagttac	tgcgacgatg	5400
ggatataagc	atgagcccct	agatcttaaa	gaagcctata	ctcaaattgc	aagacccaat	5460
ttcatgctaa	aaatcattca	aggttatgat	ggtaagccaa	gaatttgtga	actcatctgt	5520
gcagaaaata	ctgatataac	tatccacggt	gcttggaactg	gaagtgcacg	cctacaatta	5580
tttagccatg	cactagctcc	tcttgctgat	ttacctgtat	tagagatcgt	atcagcatct	5640
catatcctaa	cagatttaac	tcttgaaca	cctaagggtg	tacatgatta	tctttcagta	5700
aaataaaagc	aatatagaat	aaccactaca	aaagtagtgg	ttattctata	ttttaaatca	5760
aactgtaaaa	cttaagtttt	atagtaccta	ataatatttt	actaccagca	ttagattagt	5820
taaaatacaa	agtttggtg	aaaagtattt	tagattgcat	aatagccttc	tatactttta	5880
acaatataac	caattgctca	ccatctgctt	agaatatgct	tctttaagct	ctaaaataca	5940
tataaaaaag	taggaatttc	ttattaaaa	tcctacttat	attatatata	aatttaatcg	6000

```

ttaggtttta ttgcattgt tcctctttaa tttatctctt ataacatttt attataattg      6060
ttcatataat taattcaata tactattata tttttcaag cattaataat tattcagcat      6120
ctgtcattac atatgcttcc atactttgac ttcttattaa atcatagcta atccatccat      6180
agccattgat tccccagtct ttaccccatg aatttattat ttttacagct tttttactat      6240
catcataacc aactacgcaa actgcatgac cacctctatt ttctccatca atctgggtcat      6300
aaattggatt atcagaattt aaattatcaa aatctggata tactgatatt ccaataacta      6360
ctggatttcc agctgctatt tgtgccttta ttgcattata gtcaccatct ggaagttgac      6420
tccaactttt tgctttatat ttggctgcat tagccttttg ttcattctgta ggtgtaacct      6480
cccaactata ttactacca tcataaggca tatcagataa tgtagtacia ccttggttctt      6540
ctaataattt aaatgcat                                                    6558

```

```

<210> 63
<211> 862
<212> PRT
<213> Clostridium acetobutylicum

```

```
<400> 63
```

```
Met Lys Val Thr Thr Val Lys Glu Leu Asp Glu Lys Leu Lys Val Ile
1          5          10          15
```

```
Lys Glu Ala Gln Lys Lys Phe Ser Cys Tyr Ser Gln Glu Met Val Asp
          20          25          30
```

```
Glu Ile Phe Arg Asn Ala Ala Met Ala Ala Ile Asp Ala Arg Ile Glu
          35          40          45
```

```
Leu Ala Lys Ala Ala Val Leu Glu Thr Gly Met Gly Leu Val Glu Asp
          50          55          60
```

```
Lys Val Ile Lys Asn His Phe Ala Gly Glu Tyr Ile Tyr Asn Lys Tyr
65          70          75          80
```

```
Lys Asp Glu Lys Thr Cys Gly Ile Ile Glu Arg Asn Glu Pro Tyr Gly
          85          90          95
```

```
Ile Thr Lys Ile Ala Glu Pro Ile Gly Val Val Ala Ala Ile Ile Pro
          100          105          110
```

```
Val Thr Asn Pro Thr Ser Thr Thr Ile Phe Lys Ser Leu Ile Ser Leu
          115          120          125
```

```
Lys Thr Arg Asn Gly Ile Phe Phe Ser Pro His Pro Arg Ala Lys Lys
          130          135          140
```

```
Ser Thr Ile Leu Ala Ala Lys Thr Ile Leu Asp Ala Ala Val Lys Ser
          145          150          155          160
```

Gly Ala Pro Glu Asn Ile Ile Gly Trp Ile Asp Glu Pro Ser Ile Glu
 165 170 175
 Leu Thr Gln Tyr Leu Met Gln Lys Ala Asp Ile Thr Leu Ala Thr Gly
 180 185 190
 Gly Pro Ser Leu Val Lys Ser Ala Tyr Ser Ser Gly Lys Pro Ala Ile
 195 200 205
 Gly Val Gly Pro Gly Asn Thr Pro Val Ile Ile Asp Glu Ser Ala His
 210 215 220
 Ile Lys Met Ala Val Ser Ser Ile Ile Leu Ser Lys Thr Tyr Asp Asn
 225 230 235 240
 Gly Val Ile Cys Ala Ser Glu Gln Ser Val Ile Val Leu Lys Ser Ile
 245 250 255
 Tyr Asn Lys Val Lys Asp Glu Phe Gln Glu Arg Gly Ala Tyr Ile Ile
 260 265 270
 Lys Lys Asn Glu Leu Asp Lys Val Arg Glu Val Ile Phe Lys Asp Gly
 275 280 285
 Ser Val Asn Pro Lys Ile Val Gly Gln Ser Ala Tyr Thr Ile Ala Ala
 290 295 300
 Met Ala Gly Ile Lys Val Pro Lys Thr Thr Arg Ile Leu Ile Gly Glu
 305 310 315 320
 Val Thr Ser Leu Gly Glu Glu Glu Pro Phe Ala His Glu Lys Leu Ser
 325 330 335
 Pro Val Leu Ala Met Tyr Glu Ala Asp Asn Phe Asp Asp Ala Leu Lys
 340 345 350
 Lys Ala Val Thr Leu Ile Asn Leu Gly Gly Leu Gly His Thr Ser Gly
 355 360 365
 Ile Tyr Ala Asp Glu Ile Lys Ala Arg Asp Lys Ile Asp Arg Phe Ser
 370 375 380
 Ser Ala Met Lys Thr Val Arg Thr Phe Val Asn Ile Pro Thr Ser Gln
 385 390 395 400
 Gly Ala Ser Gly Asp Leu Tyr Asn Phe Arg Ile Pro Pro Ser Phe Thr
 405 410 415
 Leu Gly Cys Gly Phe Trp Gly Gly Asn Ser Val Ser Glu Asn Val Gly
 420 425 430

20130927_CL5646W0PCT_ST25. txt

Pro Lys His Leu Leu Asn Ile Lys Thr Val Ala Glu Arg Arg Glu Asn
435 440 445

Met Leu Trp Phe Arg Val Pro His Lys Val Tyr Phe Lys Phe Gly Cys
450 455 460

Leu Gln Phe Ala Leu Lys Asp Leu Lys Asp Leu Lys Lys Lys Arg Ala
465 470 475 480

Phe Ile Val Thr Asp Ser Asp Pro Tyr Asn Leu Asn Tyr Val Asp Ser
485 490 495

Ile Ile Lys Ile Leu Glu His Leu Asp Ile Asp Phe Lys Val Phe Asn
500 505 510

Lys Val Gly Arg Glu Ala Asp Leu Lys Thr Ile Lys Lys Ala Thr Glu
515 520 525

Glu Met Ser Ser Phe Met Pro Asp Thr Ile Ile Ala Leu Gly Gly Thr
530 535 540

Pro Glu Met Ser Ser Ala Lys Leu Met Trp Val Leu Tyr Glu His Pro
545 550 555 560

Glu Val Lys Phe Glu Asp Leu Ala Ile Lys Phe Met Asp Ile Arg Lys
565 570 575

Arg Ile Tyr Thr Phe Pro Lys Leu Gly Lys Lys Ala Met Leu Val Ala
580 585 590

Ile Thr Thr Ser Ala Gly Ser Gly Ser Glu Val Thr Pro Phe Ala Leu
595 600 605

Val Thr Asp Asn Asn Thr Gly Asn Lys Tyr Met Leu Ala Asp Tyr Glu
610 615 620

Met Thr Pro Asn Met Ala Ile Val Asp Ala Glu Leu Met Met Lys Met
625 630 635 640

Pro Lys Gly Leu Thr Ala Tyr Ser Gly Ile Asp Ala Leu Val Asn Ser
645 650 655

Ile Glu Ala Tyr Thr Ser Val Tyr Ala Ser Glu Tyr Thr Asn Gly Leu
660 665 670

Ala Leu Glu Ala Ile Arg Leu Ile Phe Lys Tyr Leu Pro Glu Ala Tyr
675 680 685

Lys Asn Gly Arg Thr Asn Glu Lys Ala Arg Glu Lys Met Ala His Ala
690 695 700

20130927_CL5646WOPCT_ST25. txt

Ser Thr Met Ala Gly Met Ala Ser Ala Asn Ala Phe Leu Gly Leu Cys
705 710 715 720

His Ser Met Ala Ile Lys Leu Ser Ser Glu His Asn Ile Pro Ser Gly
725 730 735

Ile Ala Asn Ala Leu Leu Ile Glu Glu Val Ile Lys Phe Asn Ala Val
740 745 750

Asp Asn Pro Val Lys Gln Ala Pro Cys Pro Gln Tyr Lys Tyr Pro Asn
755 760 765

Thr Ile Phe Arg Tyr Ala Arg Ile Ala Asp Tyr Ile Lys Leu Gly Gly
770 775 780

Asn Thr Asp Glu Glu Lys Val Asp Leu Leu Ile Asn Lys Ile His Glu
785 790 795 800

Leu Lys Lys Ala Leu Asn Ile Pro Thr Ser Ile Lys Asp Ala Gly Val
805 810 815

Leu Glu Glu Asn Phe Tyr Ser Ser Leu Asp Arg Ile Ser Glu Leu Ala
820 825 830

Leu Asp Asp Gln Cys Thr Gly Ala Asn Pro Arg Phe Pro Leu Thr Ser
835 840 845

Glu Ile Lys Glu Met Tyr Ile Asn Cys Phe Lys Lys Gln Pro
850 855 860

<210> 64
<211> 1665
<212> DNA
<213> Clostridium acetobutylicum

<400> 64
ttgaagagtg aatacacaat tggaagatat ttgttagacc gtttatcaga gttgggtatt 60
cggcatatct ttggtgtacc tggagattac aatctatcct ttttagacta tataatggag 120
tacaaaggga tagattgggt tggaattgc aatgaattga atgctgggta tgctgctgat 180
ggatatgcaa gaataaatgg aattggagcc atacttaca catttggtgt tggagaatta 240
agtgccatta acgcaattgc tggggcatac gctgagcaag ttccagttgt taaaattaca 300
ggtatcccca cagcaaaagt tagggacaat ggattatatg tacaccacac attaggtgac 360
ggaaggtttg atcacttttt tgaaatgttt agagaagtaa cagttgctga ggcattacta 420
agcgaagaaa atgcagcaca agaaattgat cgtgttctta ttcatgctg gagacaaaaa 480
cgtcctgttc ttataaatTT accgattgat gtatatgata aaccaattaa caaaccatta 540
aagccattac tcgattatac tatttcaagt aacaaagagg ctgcatgtga atttgttaca 600
gaaatagtac ctataataaa tagggcaaaa aagcctgtta ttcttgaga ttatggagta 660

```

tatcgttacc aagttcaaca tgtgcttaaa aacttggccg aaaaaaccgg atttcctgtg      720
gctacactaa gtatgggaaa aggtgttttc aatgaagcac accctcaatt tattgggtgtt      780
tataatgggtg atgtaagttc tccttattta aggcagcgag ttgatgaagc agactgcatt      840
attagcgttg gtgtaaaatt gacggattca accacagggg gatcttctca tggattttct      900
aaaaggaatg taattcacat tgatcctttt tcaataaagg caaaaggtaa aaaatatgca      960
cctattacga tgaaagatgc tttacagaa ttaacaagta aaattgagca tagaaacttt     1020
gaggatttag atataaagcc ttacaaatca gataatcaaa agtattttgc aaaagagaag     1080
ccaattacac aaaaacgttt ttttgagcgt attgctcact ttataaaaga aaaagatgta     1140
ttattagcag aacagggtac atgctttttt ggtgctcaa ccatacaact acccaaagat     1200
gcaactttta ttggtaacc tttatgggga tctattggat acacacttcc tgctttatta     1260
ggttcacaat tagctgatca aaaaaggcgt aatattcttt taattgggga tgggtgcattt     1320
caaatgacag cacaagaaat ttcaacaatg cttcgtttac aaatcaaacc tattatTTTT     1380
ttaattaata acgatgggta tacaattgaa cgtgctattc atggtagaga acaagtatat     1440
aacaatattc aaatgtggcg atatcataat gttccaaagg ttttaggtcc taaagaatgc     1500
agcttaacct ttaaagtaca aagtgaaact gaacttgaaa aggctctttt agtggcagat     1560
aaggattgtg aacatttgat ttttatagaa gttgttatgg atcgttatga taaacccgag     1620
cctttagaac gtctttcgaa acgttttgca aatcaaaata attag                        1665

```

```

<210> 65
<211> 858
<212> PRT
<213> Clostridium acetobutylicum

```

```
<400> 65
```

```
Met Lys Val Thr Asn Gln Lys Glu Leu Lys Gln Lys Leu Asn Glu Leu
1          5          10         15
```

```
Arg Glu Ala Gln Lys Lys Phe Ala Thr Tyr Thr Gln Glu Gln Val Asp
          20          25         30
```

```
Lys Ile Phe Lys Gln Cys Ala Ile Ala Ala Ala Lys Glu Arg Ile Asn
          35          40         45
```

```
Leu Ala Lys Leu Ala Val Glu Glu Thr Gly Ile Gly Leu Val Glu Asp
          50          55         60
```

```
Lys Ile Ile Lys Asn His Phe Ala Ala Glu Tyr Ile Tyr Asn Lys Tyr
65          70          75         80
```

```
Lys Asn Glu Lys Thr Cys Gly Ile Ile Asp His Asp Asp Ser Leu Gly
          85          90         95
```

```
Ile Thr Lys Val Ala Glu Pro Ile Gly Ile Val Ala Ala Ile Val Pro
Page 100

```

100

105

110

Thr Thr Asn Pro Thr Ser Thr Ala Ile Phe Lys Ser Leu Ile Ser Leu
 115 120 125

Lys Thr Arg Asn Ala Ile Phe Phe Ser Pro His Pro Arg Ala Lys Lys
 130 135 140

Ser Thr Ile Ala Ala Ala Lys Leu Ile Leu Asp Ala Ala Val Lys Ala
 145 150 155 160

Gly Ala Pro Lys Asn Ile Ile Gly Trp Ile Asp Glu Pro Ser Ile Glu
 165 170 175

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

Gly Pro Ser Met Val Lys Ala Ala Tyr Ser Ser Gly Lys Pro Ala Ile
 195 200 205

Gly Val Gly Ala Gly Asn Thr Pro Ala Ile Ile Asp Glu Ser Ala Asp
 210 215 220

Ile Asp Met Ala Val Ser Ser Ile Ile Leu Ser Lys Thr Tyr Asp Asn
 225 230 235 240

Gly Val Ile Cys Ala Ser Glu Gln Ser Ile Leu Val Met Asn Ser Ile
 245 250 255

Tyr Glu Lys Val Lys Glu Glu Phe Val Lys Arg Gly Ser Tyr Ile Leu
 260 265 270

Asn Gln Asn Glu Ile Ala Lys Ile Lys Glu Thr Met Phe Lys Asn Gly
 275 280 285

Ala Ile Asn Ala Asp Ile Val Gly Lys Ser Ala Tyr Ile Ile Ala Lys
 290 295 300

Met Ala Gly Ile Glu Val Pro Gln Thr Thr Lys Ile Leu Ile Gly Glu
 305 310 315 320

Val Gln Ser Val Glu Lys Ser Glu Leu Phe Ser His Glu Lys Leu Ser
 325 330 335

Pro Val Leu Ala Met Tyr Lys Val Lys Asp Phe Asp Glu Ala Leu Lys
 340 345 350

Lys Ala Gln Arg Leu Ile Glu Leu Gly Gly Ser Gly His Thr Ser Ser
 355 360 365

Leu Tyr Ile Asp Ser Gln Asn Asn Lys Asp Lys Val Lys Glu Phe Gly
 Page 101

370

375

380

Leu Ala Met Lys Thr Ser Arg Thr Phe Ile Asn Met Pro Ser Ser Gln
 385 390 395 400

Gly Ala Ser Gly Asp Leu Tyr Asn Phe Ala Ile Ala Pro Ser Phe Thr
 405 410 415

Leu Gly Cys Gly Thr Trp Gly Gly Asn Ser Val Ser Gln Asn Val Glu
 420 425 430

Pro Lys His Leu Leu Asn Ile Lys Ser Val Ala Glu Arg Arg Glu Asn
 435 440 445

Met Leu Trp Phe Lys Val Pro Gln Lys Ile Tyr Phe Lys Tyr Gly Cys
 450 455 460

Leu Arg Phe Ala Leu Lys Glu Leu Lys Asp Met Asn Lys Lys Arg Ala
 465 470 475 480

Phe Ile Val Thr Asp Lys Asp Leu Phe Lys Leu Gly Tyr Val Asn Lys
 485 490 495

Ile Thr Lys Val Leu Asp Glu Ile Asp Ile Lys Tyr Ser Ile Phe Thr
 500 505 510

Asp Ile Lys Ser Asp Pro Thr Ile Asp Ser Val Lys Lys Gly Ala Lys
 515 520 525

Glu Met Leu Asn Phe Glu Pro Asp Thr Ile Ile Ser Ile Gly Gly Gly
 530 535 540

Ser Pro Met Asp Ala Ala Lys Val Met His Leu Leu Tyr Glu Tyr Pro
 545 550 555 560

Glu Ala Glu Ile Glu Asn Leu Ala Ile Asn Phe Met Asp Ile Arg Lys
 565 570 575

Arg Ile Cys Asn Phe Pro Lys Leu Gly Thr Lys Ala Ile Ser Val Ala
 580 585 590

Ile Pro Thr Thr Ala Gly Thr Gly Ser Glu Ala Thr Pro Phe Ala Val
 595 600 605

Ile Thr Asn Asp Glu Thr Gly Met Lys Tyr Pro Leu Thr Ser Tyr Glu
 610 615 620

Leu Thr Pro Asn Met Ala Ile Ile Asp Thr Glu Leu Met Leu Asn Met
 625 630 635 640

Pro Arg Lys Leu Thr Ala Ala Thr Gly Ile Asp Ala Leu Val His Ala
 Page 102

645

650

655

I l e G l u A l a T y r V a l S e r V a l M e t A l a T h r A s p T y r T h r A s p G l u L e u
 660 665 670

A l a L e u A r g A l a I l e L y s M e t I l e P h e L y s T y r L e u P r o A r g A l a T y r
 675 680 685

L y s A s n G l y T h r A s n A s p I l e G l u A l a A r g G l u L y s M e t A l a H i s A l a
 690 695 700

S e r A s n I l e A l a G l y M e t A l a P h e A l a A s n A l a P h e L e u G l y V a l C y s
 705 710 715 720

H i s S e r M e t A l a H i s L y s L e u G l y A l a M e t H i s H i s V a l P r o H i s G l y
 725 730 735

I l e A l a C y s A l a V a l L e u I l e G l u G l u V a l I l e L y s T y r A s n A l a T h r
 740 745 750

A s p C y s P r o T h r L y s G l n T h r A l a P h e P r o G l n T y r L y s S e r P r o A s n
 755 760 765

A l a L y s A r g L y s T y r A l a G l u I l e A l a G l u T y r L e u A s n L e u L y s G l y
 770 775 780

T h r S e r A s p T h r G l u L y s V a l T h r A l a L e u I l e G l u A l a I l e S e r L y s
 785 790 795 800

L e u L y s I l e A s p L e u S e r I l e P r o G l n A s n I l e S e r A l a A l a G l y I l e
 805 810 815

A s n L y s L y s A s p P h e T y r A s n T h r L e u A s p L y s M e t S e r G l u L e u A l a
 820 825 830

P h e A s p A s p G l n C y s T h r T h r A l a A s n P r o A r g T y r P r o L e u I l e S e r
 835 840 845

G l u L e u L y s A s p I l e T y r I l e L y s S e r P h e
 850 855

<210> 66
 <211> 2589
 <212> DNA
 <213> Clostridium acetobutylicum

<400> 66
 atgaaagtca caacagtaaa ggaattagat gaaaaactca aggttaattaa agaagctcaa 60
 aaaaaattct cttgttactc gcaagaaatg gttgatgaaa tctttagaaa tgcagcaatg 120
 gcagcaatcg acgcaaggat agagctagca aaagcagctg ttttggaac cggtatgggc 180
 ttagttgaag acaaggttat aaaaaatcat tttgcaggcg aatacatcta taacaaatat 240

20130927_CL5646WOPCT_ST25. txt

aaggatgaaa aaacctgcgg tataattgaa cgaaatgaac cctacggaat tacaaaaata	300
gcagaaccta taggagttgt agctgctata atccctgtaa caaacccac atcaacaaca	360
atatttaa at ccttaatatc ccttaaaact agaaatggaa ttttcttttc gcctcaccca	420
agggcaaaaa aatccacaat actagcagct aaaacaatac ttgatgcagc cgттаagagt	480
ggtgccccgg aaaatataat aggttgata gatgaacctt caattgaact aactcaatat	540
ttaatgcaaa aagcagatat aacccttgca actggtggtc cctcactagt taaatctgct	600
tattcttccg gaaaaccagc aataggtgtt ggtccgggta acacccagc aataattgat	660
gaatctgctc atataaaaaat ggcagtaagt tcaattatat tatccaaaac ctatgataat	720
ggtgttatat gtgcttctga acaatctgta atagtcttaa aatccatata taacaaggta	780
aaagatgagt tccaagaaag aggagcttat ataataaaga aaaacgaatt ggataaagtc	840
cgtgaagtga tttttaaaga tggatccgta aaccctaaaa tagtcggaca gtcagcttat	900
actatagcag ctatggctgg cataaaagta cctaaaacca caagaatatt aataggagaa	960
gttacctcct taggtgaaga agaacctttt gccacgaaa aactatctcc tgttttggct	1020
atgtatgagg ctgacaattt tgatgatgct ttaaaaaaag cagtaactct aataaactta	1080
ggaggcctcg gccatacctc aggaatatat gcagatgaaa taaaagcacg agataaaaata	1140
gatagattta gtagtgccat gaaaaccgta agaacctttg taaatatccc aacctcacia	1200
ggtgcaagtg gagatctata taattttaga ataccacctt ctttcacgct tggctgcgga	1260
ttttggggag gaaattctgt ttccgagaat gttggtccaa aacatctttt gaatattaaa	1320
accgtagctg aaaggagaga aaacatgctt tggtttagag ttccacataa agtatatttt	1380
aagttcgggt gtcttcaatt tgctttaaaa gatttaaaag atctaaagaa aaaaagagcc	1440
tttatagtta ctgatagtga cccctataat ttaaaactatg ttgattcaat aataaaaaata	1500
cttgagcacc tagatattga ttttaaagta ttttaataagg ttggaagaga agctgatctt	1560
aaaaccataa aaaaagcaac tgaagaaatg tcctccttta tgccagacac tataatagct	1620
ttaggtggta cccctgaaat gagctctgca aagctaattgt gggtagctata tgaacatcca	1680
gaagtaaaat ttgaagatct tgcaataaaa tttatggaca taagaaagag aatatatact	1740
ttcccaaaac tcggtaaaaa ggctatgtta gttgcaatta caacttctgc tggttccggt	1800
tctgaggtta ctcttttgc tttagtaact gacaataaca ctggaaataa gtacatgtta	1860
gcagattatg aatgacacc aatatggca attgtagatg cagaacttat gatgaaaatg	1920
ccaaagggat taaccgctta ttcaggtata gatgcactag taaatagtat agaagcatac	1980
acatccgtat atgcttcaga atacacaaac ggactagcac tagaggcaat acgattaata	2040
tttaaatatt tgcctgaggc ttacaaaaac ggaagaacca atgaaaaagc aagagagaaa	2100
atggctcacg cttcaactat ggcaggtatg gcatccgcta atgcatttct aggtctatgt	2160
cattccatgg caataaaatt aagttcagaa cacaatatc ctagtggcat tgccaatgca	2220
ttactaatag aagaagtaat aaaatttaac gcagttgata atcctgtaaa acaagcccct	2280

```

tgcccacaat ataagtatcc aaacaccata tttagatatg ctcgaattgc agattatata      2340
aagcttggag gaaataactga tgaggaaaag gtagatctct taattaacaa aatacatgaa      2400
ctaaaaaaag ctttaaatat accaacttca ataaaggatg caggtgtttt ggaggaaaac      2460
ttctattcct cccttgatag aatatctgaa cttgcactag atgatcaatg cacaggcgct      2520
aatcctagat ttcctcttac aagtgaataa aaagaaatgt atataaattg ttttaaaaaa      2580
caaccttaa                                                                    2589

```

```

<210> 67
<211> 307
<212> PRT
<213> Pseudomonas putida
<400> 67

```

```

Met Ser Lys Lys Leu Lys Ala Ala Ile Ile Gly Pro Gly Asn Ile Gly
1      5      10      15

```

```

Thr Asp Leu Val Met Lys Met Leu Arg Ser Glu Trp Ile Glu Pro Val
      20      25      30

```

```

Trp Met Val Gly Ile Asp Pro Asn Ser Asp Gly Leu Lys Arg Ala Arg
      35      40      45

```

```

Asp Phe Gly Met Lys Thr Thr Ala Glu Gly Val Asp Gly Leu Leu Pro
      50      55      60

```

```

His Val Leu Asp Asp Asp Ile Arg Ile Ala Phe Asp Ala Thr Ser Ala
      65      70      75      80

```

```

Tyr Val His Ala Glu Asn Ser Arg Lys Leu Asn Ala Leu Gly Val Leu
      85      90      95

```

```

Met Val Asp Leu Thr Pro Ala Ala Ile Gly Pro Tyr Cys Val Pro Pro
      100      105      110

```

```

Val Asn Leu Lys Gln His Val Gly Arg Leu Glu Met Asn Val Asn Met
      115      120      125

```

```

Val Thr Cys Gly Gly Gln Ala Thr Ile Pro Met Val Ala Ala Val Ser
      130      135      140

```

```

Arg Val Gln Pro Val Ala Tyr Ala Glu Ile Val Ala Thr Val Ser Ser
      145      150      155      160

```

```

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

```

```

Thr Thr Ala Gly Ala Ile Glu Gln Val Gly Gly Ala Arg Glu Gly Lys
      180      185      190

```

20130927_CL5646WOPCT_ST25. txt

Ala Ile Ile Val Ile Asn Pro Ala Glu Pro Pro Leu Met Met Arg Asp
195 200 205

Thr Ile His Cys Leu Thr Asp Ser Glu Pro Asp Gln Ala Ala Ile Thr
210 215 220

Ala Ser Val His Ala Met Ile Ala Glu Val Gln Lys Tyr Val Pro Gly
225 230 235 240

Tyr Arg Leu Lys Asn Gly Pro Val Phe Asp Gly Asn Arg Val Ser Ile
245 250 255

Phe Met Glu Val Glu Gly Leu Gly Asp Tyr Leu Pro Lys Tyr Ala Gly
260 265 270

Asn Leu Asp Ile Met Thr Ala Ala Ala Leu Arg Thr Gly Glu Met Phe
275 280 285

Ala Glu Glu Ile Ala Ala Gly Thr Ile Gln Leu Pro Arg Arg Asp Ile
290 295 300

Ala Leu Ala
305

<210> 68
<211> 2180
<212> DNA
<213> Pseudomonas putida

<400> 68
ggtagccctg gagccggtca aggccggcga cttcatgcgc gtcgagatcg gcggcatcgg 60
cagcgcctcc gtgcgcttca cctgatcgaa cagaggacaa acccatgagc aagaaactca 120
aggcggccat cataggcccc ggcaatatcg gtaccgatct ggtgatgaag atgctccgtt 180
ccgagtggat tgagccggtg tggatggtcg gcatcgaccc caactccgac ggcctcaaac 240
gcgcccgcga tttcggcatg aagaccacag ccgaaggcgt cgacggcctg ctcccgcacg 300
tgctggacga cgacatccgc atcgcccttcg acgccacctc ggcctatgtg catgccgaga 360
atagccgcaa gctcaacgcg cttggcgtgc tgatggtcga cctgaccccg gcggccatcg 420
gccctactg cgtgccgccc gtcaacctca agcagcatgt cggccgcctg gaaatgaacg 480
tcaacatggt cacctgcggc ggccaggcca ccatccccat ggtcgccgcg gtgtcccgcg 540
tgcagccggt ggcctacgcc gagatcgtcg ccaccgtctc ctgcgctcg gtcggcccgg 600
gcacgcgcaa gaacatcgac gagttcacc gcaccaccgc cggcgccatc gagcaggtcg 660
gcggcgccag ggaaggcaag gcgatcatcg tcatcaacct ggccgagccg ccgctgatga 720
tgcgcgacac catccactgc ctgaccgaca gcgagccgga ccaggctgcg atcaccgctt 780
cggttcacgc gatgatcgcc gaggtgcaga aatacgtgcc cggctaccgc ctgaagaacg 840
gcccgggtgtt cgacggcaac cgcgtgtcga tcttcatgga agtcgaaggc ctgggcgact 900

20130927_CL5646WOPCT_ST25. txt

acctgccc aa gtacgccggc aacctcgaca tcatgaccgc cgccgcgctg cgtaccggcg 960
 agatgttcgc cgaggaaatc gccgccggca ccattcaact gccgcgtcgc gacatcgcg 1020
 tggcttgagg agtagacca tgaatttgca cggcaagagc gtcatcctgc acgacatgag 1080
 cctgcgcgac ggcatgcacg ccaagcgcca ccagatcagc ctggagcaga tggtcgcggt 1140
 cgccaccggc ctcgatcaag ccggtatgcc gctgatcgag atcaccacg gcgacggcct 1200
 cggcggctcgt tcgatcaact acggcttccc ggcccacagt gacgaggagt acctgcgcgc 1260
 ggtgatcccg cagctcaagc aggccaaagt ctcggcgctg ctgctgcccg gcatcggcac 1320
 cgtcgaccac ctgaagatgg ccctggactg cggcgctcgc actattcgcg tggccaccca 1380
 ctgtaccgag gcggatgtct ccgagcagca catcggcatg gcgcgcaagc tgggggtcga 1440
 caccgtcggc ttcctgatga tggcgacat gatcagcgcc gagaaagtcc tggagcaggc 1500
 caagctgatg gaaagctatg gtgccaactg catctactgc accgactcgg ccggctacat 1560
 gctgcctgat gaagtcagcg agaaaatcgg cctcctgcgc gccgagctga acccgccac 1620
 cgaagtcggc ttccacggcc accacaacat gggcatggct atcgccaact cgctggccgc 1680
 catcgaagcc ggtgccgcgc gcatcgacgg ctcggtcgcc ggcctcggcg ccggtgccgg 1740
 caacaccccg ctggaagtgt tcgtcgcagt gtgcaaagc atgggcgtgg agaccggcat 1800
 cgacctgtac aagatcatgg acgtggccga ggacctggtg gtgccgatga tggatcagcc 1860
 gatccgcgtc gaccgcgacg ccctgaccct gggctacgcc ggggtgtaca gctcgttct 1920
 gctgttcgcc cagcgcgccg agaagaaata tggcgtgtcg gccgcgcaca tcctggtcga 1980
 actgggcccgg cgcggcaccg tcggtggcca ggaagacatg atcgaagacc tcgccctgga 2040
 catggcccgg gcccgtcagc agcagaaggt gagcgcatga accgtaccct gaccgcgaa 2100
 caggtgctgg ccctggccga gcacatcgaa aacgccgagc tgaatgtcca cgacatcggc 2160
 aagtgacca acgattttcc 2180

<210> 69
 <211> 307
 <212> PRT
 <213> Thermus thermophilus

<400> 69

Met Ser Glu Arg Val Lys Val Ala Ile Leu Gly Ser Gly Asn Ile Gly
 1 5 10 15

Thr Asp Leu Met Tyr Lys Leu Leu Lys Asn Pro Gly His Met Glu Leu
 20 25 30

Val Ala Val Val Gly Ile Asp Pro Lys Ser Glu Gly Leu Ala Arg Ala
 35 40 45

Arg Ala Leu Gly Leu Glu Ala Ser His Glu Gly Ile Ala Tyr Ile Leu
 50 55 60

20130927_CL5646WOPCT_ST25. txt

Glu Arg Pro Glu Ile Lys Ile Val Phe Asp Ala Thr Ser Ala Lys Ala
65 70 75 80

His Val Arg His Ala Lys Leu Leu Arg Glu Ala Gly Lys Ile Ala Ile
85 90 95

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

Leu Lys Glu His Leu Asp Lys Asp Asn Val Asn Leu Ile Thr Cys Gly
115 120 125

Gly Gln Ala Thr Ile Pro Leu Val Tyr Ala Val His Arg Val Ala Pro
130 135 140

Val Leu Tyr Ala Glu Met Val Ser Thr Val Ala Ser Arg Ser Ala Gly
145 150 155 160

Pro Gly Thr Arg Gln Asn Ile Asp Glu Phe Thr Phe Thr Thr Ala Arg
165 170 175

Gly Leu Glu Ala Ile Gly Gly Ala Lys Lys Gly Lys Ala Ile Ile Ile
180 185 190

Leu Asn Pro Ala Glu Pro Pro Ile Leu Met Thr Asn Thr Val Arg Cys
195 200 205

Ile Pro Glu Asp Glu Gly Phe Asp Arg Glu Ala Val Val Ala Ser Val
210 215 220

Arg Ala Met Glu Arg Glu Val Gln Ala Tyr Val Pro Gly Tyr Arg Leu
225 230 235 240

Lys Ala Asp Pro Val Phe Glu Arg Leu Pro Thr Pro Trp Gly Glu Arg
245 250 255

Thr Val Val Ser Met Leu Leu Glu Val Glu Gly Ala Gly Asp Tyr Leu
260 265 270

Pro Lys Tyr Ala Gly Asn Leu Asp Ile Met Thr Ala Ser Ala Arg Arg
275 280 285

Val Gly Glu Val Phe Ala Gln His Leu Leu Gly Lys Pro Val Glu Glu
290 295 300

Val Val Ala
305

<210> 70
<211> 924
<212> DNA

<213> Thermus thermophilus

<400> 70

```

atgtccgaaa gggttaaggt agccatcctg ggctccggca acatcgggac ggacctgatg      60
tacaagctcc tgaagaaccc gggccacatg gagcttgtgg cgggtggtggg gatagacccc      120
aagtccgagg gcctggcccc ggcgcgggcc ttagggttag aggcgagcca cgaagggatc      180
gcctacatcc tggagaggcc ggagatcaag atcgtctttg acgccaccag cgccaaggcc      240
cacgtgcgcc acgccaagct cctgagggag gcggggaaga tcgcataga cctcacgccg      300
gcggcccggg gcccttacgt ggtgcccccg gtgaacctga aggaacacct ggacaaggac      360
aacgtgaacc tcatcacctg cggggggcag gccaccatcc ccctggtcta cgcggtgcac      420
cgggtggccc ccgtgctcta cgcgagatg gtctccacgg tggcctcccg ctccgcgggc      480
cccggcaccg ggagaacat cgacgagttc acctcacca ccgccgggg cctggaggcc      540
atcggggggg ccaagaaggg gaaggccatc atcatcctga acccggcgga accccccatc      600
ctcatgacca acaccgtgcg ctgcatcccc gaggacgagg gctttgaccg ggaggccgtg      660
gtggcgagcg tccgggcat ggagcgggag gtccaggcct acgtgcccg ctaccgcctg      720
aaggcggacc cgggtgttga gaggttccc acccctggg gggagcgcac cgtggtctcc      780
atgctcctgg aggtggaggg ggcgggggac tatttgccca aatacgccg caacctggac      840
atcatgacgg cttctgcccg gaggggtggg gaggtcttcg ccagcacct cctggggaag      900
ccggtggagg aggtggtggc gtga                                          924

```

<210> 71

<211> 417

<212> PRT

<213> Escherichia coli

<400> 71

```

Met Thr Phe Ser Leu Phe Gly Asp Lys Phe Thr Arg His Ser Gly Ile
1          5          10          15

Thr Leu Leu Met Glu Asp Leu Asn Asp Gly Leu Arg Thr Pro Gly Ala
20          25          30

Ile Met Leu Gly Gly Gly Asn Pro Ala Gln Ile Pro Glu Met Gln Asp
35          40          45

Tyr Phe Gln Thr Leu Leu Thr Asp Met Leu Glu Ser Gly Lys Ala Thr
50          55          60

Asp Ala Leu Cys Asn Tyr Asp Gly Pro Gln Gly Lys Thr Glu Leu Leu
65          70          75          80

Thr Leu Leu Ala Gly Met Leu Arg Glu Lys Leu Gly Trp Asp Ile Glu
85          90          95

Pro Gln Asn Ile Ala Leu Thr Asn Gly Ser Gln Ser Ala Phe Phe Tyr

```

100

105

110

Leu Phe Asn Leu Phe Ala Gly Arg Arg Ala Asp Gly Arg Val Lys Lys
 115 120 125

Val Leu Phe Pro Leu Ala Pro Glu Tyr Ile Gly Tyr Ala Asp Ala Gly
 130 135 140

Leu Glu Glu Asp Leu Phe Val Ser Ala Arg Pro Asn Ile Glu Leu Leu
 145 150 155 160

Pro Glu Gly Gln Phe Lys Tyr His Val Asp Phe Glu His Leu His Ile
 165 170 175

Gly Glu Glu Thr Gly Met Ile Cys Val Ser Arg Pro Thr Asn Pro Thr
 180 185 190

Gly Asn Val Ile Thr Asp Glu Glu Leu Leu Lys Leu Asp Ala Leu Ala
 195 200 205

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

Phe Pro Gly Ile Ile Phe Ser Glu Ala Arg Pro Leu Trp Asn Pro Asn
 225 230 235 240

Ile Val Leu Cys Met Ser Leu Ser Lys Leu Gly Leu Pro Gly Ser Arg
 245 250 255

Cys Gly Ile Ile Ile Ala Asn Glu Lys Ile Ile Thr Ala Ile Thr Asn
 260 265 270

Met Asn Gly Ile Ile Ser Leu Ala Pro Gly Gly Ile Gly Pro Ala Met
 275 280 285

Met Cys Glu Met Ile Lys Arg Asn Asp Leu Leu Arg Leu Ser Glu Thr
 290 295 300

Val Ile Lys Pro Phe Tyr Tyr Gln Arg Val Gln Glu Thr Ile Ala Ile
 305 310 315 320

Ile Arg Arg Tyr Leu Pro Glu Asn Arg Cys Leu Ile His Lys Pro Glu
 325 330 335

Gly Ala Ile Phe Leu Trp Leu Trp Phe Lys Asp Leu Pro Ile Thr Thr
 340 345 350

Lys Gln Leu Tyr Gln Arg Leu Lys Ala Arg Gly Val Leu Met Val Pro
 355 360 365

Gly His Asn Phe Phe Pro Gly Leu Asp Lys Pro Trp Pro His Thr His
 Page 110

370

Gln Cys Met Arg Met Asn Tyr Val Pro Glu Pro Glu Lys Ile Glu Ala
385 390 395 400

Gly Val Lys Ile Leu Ala Glu Glu Ile Glu Arg Ala Trp Ala Glu Ser
405 410 415

His

<210> 72
<211> 417
<212> PRT
<213> Escherichia coli
<400> 72

Met Thr Phe Ser Leu Phe Gly Asp Lys Phe Thr Arg His Ser Gly Ile
1 5 10 15

Thr Leu Leu Met Glu Asp Leu Asn Asp Gly Leu Arg Thr Pro Gly Ala
20 25 30

Ile Met Leu Gly Gly Gly Asn Pro Ala Gln Ile Pro Glu Met Gln Asp
35 40 45

Tyr Phe Gln Thr Leu Leu Thr Asp Met Leu Glu Ser Gly Lys Ala Thr
50 55 60

Asp Ala Leu Cys Asn Tyr Asp Gly Pro Gln Gly Lys Thr Glu Leu Leu
65 70 75 80

Thr Leu Leu Ala Gly Met Leu Arg Glu Lys Leu Gly Trp Asp Ile Glu
85 90 95

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

Leu Phe Asn Leu Phe Ala Gly Arg Arg Ala Asp Gly Arg Val Lys Lys
115 120 125

Val Leu Phe Pro Leu Ala Pro Glu Tyr Ile Gly Tyr Ala Asp Ala Gly
130 135 140

Leu Glu Glu Asp Leu Phe Val Ser Ala Arg Pro Asn Ile Glu Leu Leu
145 150 155 160

Pro Glu Gly Gln Phe Lys Tyr His Val Asp Phe Glu His Leu His Ile
165 170 175

Gly Glu Glu Thr Gly Met Ile Cys Val Ser Arg Pro Thr Asn Pro Thr
180 185 190

Gly Asn Val Ile Thr Asp Glu Glu Leu Leu Lys Leu Asp Ala Leu Ala
195 200 205

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

Phe Pro Gly Ile Ile Phe Ser Glu Ala Arg Pro Leu Trp Asn Pro Asn
225 230 235 240

Ile Val Leu Cys Met Ser Leu Ser Lys Leu Gly Leu Pro Gly Ser Arg
245 250 255

Cys Gly Ile Ile Ile Ala Asn Glu Lys Ile Ile Thr Ala Ile Thr Asn
260 265 270

Met Asn Gly Ile Ile Ser Leu Ala Pro Gly Gly Ile Gly Pro Ala Met
275 280 285

Met Cys Glu Met Ile Lys Arg Asn Asp Leu Leu Arg Leu Ser Glu Thr
290 295 300

Val Ile Lys Pro Phe Tyr Tyr Gln Arg Val Gln Glu Thr Ile Ala Ile
305 310 315 320

Ile Arg Arg Tyr Leu Pro Glu Asn Arg Cys Leu Ile His Lys Pro Glu
325 330 335

Gly Ala Ile Phe Leu Trp Leu Trp Phe Lys Asp Leu Pro Ile Thr Thr
340 345 350

Lys Gln Leu Tyr Gln Arg Leu Lys Ala Arg Gly Val Leu Met Val Pro
355 360 365

Gly His Asn Phe Phe Pro Gly Leu Asp Lys Pro Trp Pro His Thr His
370 375 380

Gln Cys Met Arg Met Asn Tyr Val Pro Glu Pro Glu Lys Ile Glu Ala
385 390 395 400

Gly Val Lys Ile Leu Ala Glu Glu Ile Glu Arg Ala Trp Ala Glu Ser
405 410 415

His

<210> 73
<211> 425
<212> PRT
<213> Bacillus licheniformis

<400> 73

20130927_CL5646WOPCT_ST25. txt

Met Lys Pro Pro Leu Ser Lys Ile Gly Glu Lys Met Ile Glu Lys Thr
1 5 10 15

Gly Val Arg Ala Val Met Ser Asp Ile Gln Glu Val Leu Ala Gly Gly
20 25 30

Glu Arg Ser Tyr Ile Asn Leu Ser Ala Gly Asn Pro Met Ile Leu Pro
35 40 45

Gly Val Ser Ala Met Trp Lys Ser Ala Leu Ala Asp Leu Leu Asp Asp
50 55 60

Asp Arg Phe Ser Ser Val Ile Gly Gln Tyr Gly Ser Ser Tyr Gly Thr
65 70 75 80

Asp Glu Leu Ile Ala Ser Val Val Arg Phe Phe Ser Glu Arg Tyr Ser
85 90 95

Ala Gly Ile Arg Lys Glu Asn Val Leu Ile Thr Ala Gly Ser Gln Gln
100 105 110

Leu Phe Phe Leu Ala Ile Asn Ser Phe Cys Gly Met Gly Ser Gly Ser
115 120 125

Val Met Lys Lys Ala Leu Ile Pro Met Leu Pro Asp Tyr Ser Gly Tyr
130 135 140

Ser Gly Ala Ala Leu Glu Arg Glu Met Ile Glu Gly Ile Pro Pro Leu
145 150 155 160

Ile Ser Lys Leu Asp Asp His Thr Phe Arg Tyr Glu Leu Asp Arg Lys
165 170 175

Gly Phe Leu Glu Arg Met Arg Ile Gly Ala Val Leu Leu Ser Arg Pro
180 185 190

Asn Asn Pro Cys Gly Asn Ile Leu Pro Lys Glu Asp Val Ala Phe Ile
195 200 205

Ser Asp Ala Cys Arg Glu Ala Asn Val Pro Leu Phe Ile Asp Ser Ala
210 215 220

Tyr Ala Pro Pro Phe Pro Ala Ile His Phe Ile Asp Met Glu Pro Ile
225 230 235 240

Phe Asn Glu Gln Ile Ile His Cys Met Ser Leu Ser Lys Ala Gly Leu
245 250 255

Pro Gly Glu Arg Ile Gly Ile Ala Ile Gly Pro Ser Arg Tyr Ile Gln
260 265 270

20130927_CL5646WOPCT_ST25. txt

Ala Met Glu Ala Phe Gln Ser Asn Ala Ala Ile His Ser Ser Arg Leu
275 280 285

Gly Gln Tyr Met Ala Ala Ser Val Leu Asn Asp Gly Arg Leu Ala Asp
290 295 300

Val Ser Leu Asn Glu Val Arg Pro Tyr Tyr Arg Asn Lys Phe Met Leu
305 310 315 320

Leu Lys Glu Thr Leu Leu Cys Lys Met Pro Glu Asp Ile Lys Trp Tyr
325 330 335

Leu His Gln Gly Glu Gly Ser Leu Phe Gly Trp Leu Trp Phe Glu Asp
340 345 350

Leu Pro Val Thr Asp Ala Ala Leu Tyr Glu Tyr Met Lys Ala Asp Gly
355 360 365

Val Ile Ile Val Pro Gly Ser Ser Phe Phe His Arg Gln Ser Arg Arg
370 375 380

Leu Ala His Ser His Gln Cys Ile Arg Ile Ser Leu Thr Ala Ala Asp
385 390 395 400

Glu Asp Ile Ile Arg Gly Ile Asp Val Leu Ala Lys Ile Ala Lys Gly
405 410 415

Val Tyr Glu Lys Gln Val Glu Tyr Leu
420 425

<210> 74
<211> 1278
<212> DNA
<213> Bacillus licheniformis

<400> 74
ttataagtat tcaacctgtt tctcatatac acccttcgca attttagcta aaacatcgat 60
tccccttata atatcttcat ccgccgcggt taggctgatt cgtatacact ggtgtgaatg 120
cgccaggcgc cgggattgac ggtgaaagaa agatgatccg ggaacgataa tgactccatc 180
cgctttcata tactcataca gcgctgcatc ggtcaccggc aggtcttcaa accacagcca 240
tccgaaaagc gatccttccc cttgatgcag ataccatttg atgtcttcag gcatcttgca 300
taaaagcggt tccttgagca gcatgaattt attgcggtaa tatggcctga cttcattcag 360
cgacacgtcg gcgaggcgcc cgtcattcaa tactgatgca gccatatact gccccagcct 420
tgaagaatgg atcgccgcat tcgactgaaa agcttccatt gcctgaatat accgggacgg 480
cccgatggcg attccgatcc tttcgccagg caggccggct tttgaaaggc tcatcacagt 540
aatgatctgc tcgttgaaaa tcggttccat gtcgataaag tgaatcgccg gaaaaggcgg 600
agcatatgcg gaatcaatga acagcggaac attcgcttct cggcatgcgt ctgaaatgaa 660

20130927_CL5646WOPCT_ST25.txt

tgctacatct tctttaggca agatgtttcc gcaaggattg ttcgggcgcg atagcaagac 720
 agcaccgatg cgcacacctt ctaaaaaccc cttacggtcg agtcatatc gaaacgtatg 780
 atcatccaat ttcgatatga gcggagggat cccctcaatc atctcccgct ccagtgccgc 840
 cccgctgtat cccgaatagt caggcagcat cgggatcaag gcttttttca tcacagatcc 900
 gcttccatt cgcgaaaacg aattgatcgc cagaaaaaac agctgctggc ttccggctgt 960
 aatcaacacg ttctcttttc gaatgccggc gctataccgc tctgaaaaga agcggacaac 1020
 acttgcaatc agttcatcgg ttccatagct cgatccgtat tggccgatca ccgaagaaaa 1080
 cctgtcatcg tcaaggagat cggcaagagc cgacttccac atggctgaca cgccgggcaa 1140
 aatcatcggg ttgcccgcac ttaaattaat gtatgaccgt tcaccgccgg ccaggacttc 1200
 ctgaatatcg ctcatcacag ccctgacccc tgtttttca atcattttct ctccgatttt 1260
 gcttaatggc ggcttcac 1278

<210> 75
 <211> 309
 <212> PRT
 <213> Escherichia coli

<400> 75

Met Thr Thr Lys Lys Ala Asp Tyr Ile Trp Phe Asn Gly Glu Met Val
 1 5 10 15
 Arg Trp Glu Asp Ala Lys Val His Val Met Ser His Ala Leu His Tyr
 20 25 30
 Gly Thr Ser Val Phe Glu Gly Ile Arg Cys Tyr Asp Ser His Lys Gly
 35 40 45
 Pro Val Val Phe Arg His Arg Glu His Met Gln Arg Leu His Asp Ser
 50 55 60
 Ala Lys Ile Tyr Arg Phe Pro Val Ser Gln Ser Ile Asp Glu Leu Met
 65 70 75 80
 Glu Ala Cys Arg Asp Val Ile Arg Lys Asn Asn Leu Thr Ser Ala Tyr
 85 90 95
 Ile Arg Pro Leu Ile Phe Val Gly Asp Val Gly Met Gly Val Asn Pro
 100 105 110
 Pro Ala Gly Tyr Ser Thr Asp Val Ile Ile Ala Ala Phe Pro Trp Gly
 115 120 125
 Ala Tyr Leu Gly Ala Glu Ala Leu Glu Gln Gly Ile Asp Ala Met Val
 130 135 140
 Ser Ser Trp Asn Arg Ala Ala Pro Asn Thr Ile Pro Thr Ala Ala Lys

145 150 — 155 — 160

Ala Gly Gly Asn Tyr Leu Ser Ser Leu Leu Val Gly Ser Glu Ala Arg
165 170 175

Arg His Gly Tyr Gln Glu Gly Ile Ala Leu Asp Val Asn Gly Tyr Ile
180 185 190

Ser Gl u Gly Ala Gly Gl u Asn Leu Phe Gl u Val Lys Asp Gly Val Leu
195 200 205

Phe Thr Pro Pro Phe Thr Ser Ser Ala Leu Pro Gly Ile Thr Arg Asp
210 215 220

Ala Ile Ile Lys Leu Ala Lys Glu Leu Gly Ile Glu Val Arg Glu Gln
225 230 235 240

Val Leu Ser Arg Gl u Ser Leu Tyr Leu Al a Asp Gl u Val Phe Met Ser
245 250 255

Gly Thr Ala Ala Glu Ile Thr Pro Val Arg Ser Val Asp Gly Ile Gln
260 265 270

Val Gly Glu Gly Arg Cys Gly Pro Val Thr Lys Arg Ile Gln Gln Ala
275 280 285

Phe Phe Gly Leu Phe Thr Gly Glu Thr Glu Asp Lys Trp Gly Trp Leu
290 295 300

Asp Gl n Val Asn Gl n
305

<210>	76
<211>	1476
<212>	DNA
<213>	Escheri chi a col i

<400>	76								
atggctaact	acttcaatac	actgaatctg	cgccagcagc	tggcacagct	gggcaaagt				60
cgctttatgg	gccgcgatga	attcgccgat	ggcgcgagct	accttcagg	taaaaaagta				120
gtcatcgtcg	gctgtggcgc	acaggggtctg	aaccagggcc	tgaacatgcg	tgattctggt				180
ctcgatatct	cctacgctct	gcgtaaagaa	gcgattgccg	agaagcgcgc	gtcctggcgt				240
aaagcgaccg	aaaatggttt	taaagtgggt	acttacgaag	aactgatccc	acaggcggat				300
ctgggtgatta	acctgacgcc	ggacaagcag	cactctgatg	tagtgcgcac	cgtacagcca				360
ctgatgaaag	acggcgcggc	gctgggctac	tcgcacggtt	tcaacatcgt	cgaagtgggc				420
gagcagatcc	gtaaagatat	caccgtagtg	atggttgcg	cgaaatgccc	aggcaccgaa				480
gtgcgtgaag	agtacaaacg	tgggttcggc	gtaccgacgc	tgattgccgt	tcacccggaa				540
aacgatccga	aaggcgaagg	catggcgatt	gccaaagcct	gggcggctgc	aaccggtggt				600

```

caccgtgcgg gtgtgctgga atcgtccttc gttgcggaag tgaaatctga cctgatgggc      660
gagcaaacca tcctgtgcgg tatgttgtag gctggctctc tgctgtgctt cgacaagctg      720
gtggaagaag gtaccgatcc agcatatcga gaaaaactga ttcagttcgg ttgggaaacc      780
atcaccgaag cactgaaaca gggcggcatc accctgatga tggaccgtct ctctaaccgc      840
gcgaaactgc gtgcttatgc gctttctgaa cagctgaaag agatcatggc acccctgttc      900
cagaaacata tggacgacat catctccggc gaattctctt ccggtatgat ggcggactgg      960
gccaacgatg ataagaaact gctgacctgg cgtgaagaga ccggcaaac cgcgtttgaa     1020
accgcgccgc agtatgaagg caaaatcggc gagcaggagt acttcgataa aggcgtactg     1080
atgattgcga tggtgaaagc gggcggtgaa ctggcgttcg aaaccatggt cgattccggc     1140
atcattgaag agtctgcata ttatgaatca ctgcacgagc tgccgctgat tgccaacacc     1200
atcgcccgtg agcgtctgta cgaaatgaac gtggttatct ctgataccgc tgagtacggt     1260
aactatctgt tctcttacgc ttgtgtgccg ttgctgaaac cgtttatggc agagctgcaa     1320
ccgggcgacc tgggtaaagc tattccgga ggcgcggtag ataacgggca actgcgtgat     1380
gtgaacgaag cgattcgtag ccatgcgatt gagcaggtag gtaagaaact gcgcggctat     1440
atgacagata tgaaacgtat tgctgttgcg ggttaa                                1476

```

```

<210> 77
<211> 376
<212> PRT
<213> Saccharomyces cerevisiae
<400> 77

```

```

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

```

```

Gln His Ala Ser Lys Pro Lys Pro Asn Ser Glu Leu Val Phe Gly Lys
20          25          30

```

```

Ser Phe Thr Asp His Met Leu Thr Ala Glu Trp Thr Ala Glu Lys Gly
35          40          45

```

```

Trp Gly Thr Pro Glu Ile Lys Pro Tyr Gln Asn Leu Ser Leu Asp Pro
50          55          60

```

```

Ser Ala Val Val Phe His Tyr Ala Phe Glu Leu Phe Glu Gly Met Lys
65          70          75          80

```

```

Ala Tyr Arg Thr Val Asp Asn Lys Ile Thr Met Phe Arg Pro Asp Met
85          90          95

```

```

Asn Met Lys Arg Met Asn Lys Ser Ala Gln Arg Ile Cys Leu Pro Thr
100         105         110

```

```

Phe Asp Pro Glu Glu Leu Ile Thr Leu Ile Gly Lys Leu Ile Gln Gln
Page 117

```

115

120

125

Asp Lys Cys Leu Val Pro Glu Gly Lys Gly Tyr Ser Leu Tyr Ile Arg
 130 135 140
 Pro Thr Leu Ile Gly Thr Thr Ala Gly Leu Gly Val Ser Thr Pro Asp
 145 150 155 160
 Arg Ala Leu Leu Tyr Val Ile Cys Cys Pro Val Gly Pro Tyr Tyr Lys
 165 170 175
 Thr Gly Phe Lys Ala Val Arg Leu Glu Ala Thr Asp Tyr Ala Thr Arg
 180 185 190
 Ala Trp Pro Gly Gly Cys Gly Asp Lys Lys Leu Gly Ala Asn Tyr Ala
 195 200 205
 Pro Cys Val Leu Pro Gln Leu Gln Ala Ala Ser Arg Gly Tyr Gln Gln
 210 215 220
 Asn Leu Trp Leu Phe Gly Pro Asn Asn Asn Ile Thr Glu Val Gly Thr
 225 230 235 240
 Met Asn Ala Phe Phe Val Phe Lys Asp Ser Lys Thr Gly Lys Lys Glu
 245 250 255
 Leu Val Thr Ala Pro Leu Asp Gly Thr Ile Leu Glu Gly Val Thr Arg
 260 265 270
 Asp Ser Ile Leu Asn Leu Ala Lys Glu Arg Leu Glu Pro Ser Glu Trp
 275 280 285
 Thr Ile Ser Glu Arg Tyr Phe Thr Ile Gly Glu Val Thr Glu Arg Ser
 290 295 300
 Lys Asn Gly Glu Leu Leu Glu Ala Phe Gly Ser Gly Thr Ala Ala Ile
 305 310 315 320
 Val Ser Pro Ile Lys Glu Ile Gly Trp Lys Gly Glu Gln Ile Asn Ile
 325 330 335
 Pro Leu Leu Pro Gly Glu Gln Thr Gly Pro Leu Ala Lys Glu Val Ala
 340 345 350
 Gln Trp Ile Asn Gly Ile Gln Tyr Gly Glu Thr Glu His Gly Asn Trp
 355 360 365
 Ser Arg Val Val Thr Asp Leu Asn
 370 375

<210> 78

<211> 376
 <212> PRT
 <213> Saccharomyces cerevisiae

<400> 78

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

Gln His Ala Ser Lys Pro Lys Pro Asn Ser Glu Leu Val Phe Gly Lys
 20 25 30

Ser Phe Thr Asp His Met Leu Thr Ala Glu Trp Thr Ala Glu Lys Gly
 35 40 45

Trp Gly Thr Pro Glu Ile Lys Pro Tyr Gln Asn Leu Ser Leu Asp Pro
 50 55 60

Ser Ala Val Val Phe His Tyr Ala Phe Glu Leu Phe Glu Gly Met Lys
 65 70 75 80

Ala Tyr Arg Thr Val Asp Asn Lys Ile Thr Met Phe Arg Pro Asp Met
 85 90 95

Asn Met Lys Arg Met Asn Lys Ser Ala Gln Arg Ile Cys Leu Pro Thr
 100 105 110

Phe Asp Pro Glu Glu Leu Ile Thr Leu Ile Gly Lys Leu Ile Gln Gln
 115 120 125

Asp Lys Cys Leu Val Pro Glu Gly Lys Gly Tyr Ser Leu Tyr Ile Arg
 130 135 140

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

Arg Ala Leu Leu Tyr Val Ile Cys Cys Pro Val Gly Pro Tyr Tyr Lys
 165 170 175

Thr Gly Phe Lys Ala Val Arg Leu Glu Ala Thr Asp Tyr Ala Thr Arg
 180 185 190

Ala Trp Pro Gly Gly Cys Gly Asp Lys Lys Leu Gly Ala Asn Tyr Ala
 195 200 205

Pro Cys Val Leu Pro Gln Leu Gln Ala Ala Ser Arg Gly Tyr Gln Gln
 210 215 220

Asn Leu Trp Leu Phe Gly Pro Asn Asn Asn Ile Thr Glu Val Gly Thr
 225 230 235 240

Met Asn Ala Phe Phe Val Phe Lys Asp Ser Lys Thr Gly Lys Lys Glu
 245 250 255

Leu Val Thr Ala Pro Leu Asp Gly Thr Ile Leu Glu Gly Val Thr Arg
260 265 270

Asp Ser Ile Leu Asn Leu Ala Lys Glu Arg Leu Glu Pro Ser Glu Trp
275 280 285

Thr Ile Ser Glu Arg Tyr Phe Thr Ile Gly Glu Val Thr Glu Arg Ser
290 295 300

Lys Asn Gly Glu Leu Leu Glu Ala Phe Gly Ser Gly Thr Ala Ala Ile
305 310 315 320

Val Ser Pro Ile Lys Glu Ile Gly Trp Lys Gly Glu Gln Ile Asn Ile
325 330 335

Pro Leu Leu Pro Gly Glu Gln Thr Gly Pro Leu Ala Lys Glu Val Ala
340 345 350

Gln Trp Ile Asn Gly Ile Gln Tyr Gly Glu Thr Glu His Gly Asn Trp
355 360 365

Ser Arg Val Val Thr Asp Leu Asn
370 375

<210> 79
<211> 330
<212> PRT
<213> Methanobacterium thermoautotrophicum

<400> 79

Met Arg Leu Trp Arg Ala Leu Tyr Arg Pro Pro Thr Ile Thr Tyr Pro
1 5 10 15

Ser Lys Ser Pro Glu Val Ile Ile Met Ser Cys Glu Ala Ser Gly Lys
20 25 30

Ile Trp Leu Asn Gly Glu Met Val Glu Trp Glu Glu Ala Thr Val His
35 40 45

Val Leu Ser His Val Val His Tyr Gly Ser Ser Val Phe Glu Gly Ile
50 55 60

Arg Cys Tyr Arg Asn Ser Lys Gly Ser Ala Ile Phe Arg Leu Arg Glu
65 70 75 80

His Val Lys Arg Leu Phe Asp Ser Ala Lys Ile Tyr Arg Met Asp Ile
85 90 95

Pro Tyr Thr Gln Glu Gln Ile Cys Asp Ala Ile Val Glu Thr Val Arg
100 105 110

20130927_CL5646WOPCT_ST25. txt

Glu Asn Gly Leu Glu Glu Cys Tyr Ile Arg Pro Val Val Phe Arg Gly
115 120 125

Tyr Gly Glu Met Gly Val His Pro Val Asn Cys Pro Val Asp Val Ala
130 135 140

Val Ala Ala Trp Glu Trp Gly Ala Tyr Leu Gly Ala Glu Ala Leu Glu
145 150 155 160

Val Gly Val Asp Ala Gly Val Ser Thr Trp Arg Arg Met Ala Pro Asn
165 170 175

Thr Met Pro Asn Met Ala Lys Ala Gly Gly Asn Tyr Leu Asn Ser Gln
180 185 190

Leu Ala Lys Met Glu Ala Val Arg His Gly Tyr Asp Glu Ala Ile Met
195 200 205

Leu Asp Tyr His Gly Tyr Ile Ser Glu Gly Ser Gly Glu Asn Ile Phe
210 215 220

Leu Val Ser Glu Gly Glu Ile Tyr Thr Pro Pro Val Ser Ser Ser Leu
225 230 235 240

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

Gly Val Thr Val His Glu Glu Pro Ile Thr Arg Glu Met Leu Tyr Ile
260 265 270

Ala Asp Glu Ala Phe Phe Thr Gly Thr Ala Ala Glu Ile Thr Pro Ile
275 280 285

Arg Ser Val Asp Gly Ile Glu Ile Gly Ala Gly Arg Arg Gly Pro Val
290 295 300

Thr Lys Leu Leu Gln Asp Glu Phe Phe Arg Ile Ile Arg Ala Glu Thr
305 310 315 320

Glu Asp Ser Phe Gly Trp Leu Thr Tyr Ile
325 330

<210> 80
<211> 993
<212> DNA
<213> Methanobacterium thermoautotrophicum

<400> 80
tcagatgtag gtgagccatc cgaagctgtc ctctgtctct gccctgatta tcctgaagaa 60
ctcatcctgc agcagctttg taacgggacc ccttcgcccg gcacctatct ctataccatc 120
aactgatctg atgggtgtta tctctgcggc tgtacctgtg aagaaggcct catctgcgat 180

```

gtagagcatc tccctgggta tgggttcctc atgcacggta acaccctcgg tcctggctat      240
ctttattacg gagtcccttg ttatccccct cagaagggat gatgaaacag ggggggtgta      300
aatttcaccc tctactgacga ggaatatgtt ctccccgcta ccctcactta ttagccatg      360
gtagtccagc attatggcct catcatagcc gtgtctcaca gcctccatct tggcaagctg      420
tgagttgagg tagttaccgc cggcctttgc catgttgggc attgtgtttg gtgccatcct      480
ccgccagggt gaaacaccag catcgacacc aacctcaagg gcctctgcac ccagataggg      540
ccccattcc caggcagcca cagcgacgtc cactgggcag ttcaccgggt gaacacccat      600
ctcaccgtat cccctgaata ccacgggtct tatatagcac tcctcaagtc cgttctccct      660
gacggtctca actatggcat cacatatctg ctctgggtg tagggtatgt ccatccggta      720
tatctttgca gaatcaaaaa ggcgtttaac atgctccgc aaacggaaga tggctgaccc      780
cttactgttc ctgtagcacc ttattccctc aaagacagat gatccataat gcacaacatg      840
tgagagtacg tggacggtgg cttcttccca ttcaaccatt tcaccgttta accatatctt      900
tccactggct tcgcatgaca tgataataac ctcaggtgat ttactaggat aggttatggt      960
tggaggccta tataatgctc tccataaccg caa                                  993

```

```

<210> 81
<211> 364
<212> PRT
<213> Streptomyces coelicolor
<400> 81

```

```

Met Thr Asp Val Asn Gly Ala Pro Ala Asp Val Leu His Thr Leu Phe
1          5          10          15

His Ser Asp Gln Gly Gly His Glu Gln Val Val Leu Cys Gln Asp Arg
20          25          30

Ala Ser Gly Leu Lys Ala Val Ile Ala Leu His Ser Thr Ala Leu Gly
35          40          45

Pro Ala Leu Gly Gly Thr Arg Phe Tyr Pro Tyr Ala Ser Glu Ala Glu
50          55          60

Ala Val Ala Asp Ala Leu Asn Leu Ala Arg Gly Met Ser Tyr Lys Asn
65          70          75          80

Ala Met Ala Gly Leu Asp His Gly Gly Gly Lys Ala Val Ile Ile Gly
85          90          95

Asp Pro Glu Gln Ile Lys Ser Glu Glu Leu Leu Leu Ala Tyr Gly Arg
100         105         110

Phe Val Ala Ser Leu Gly Gly Arg Tyr Val Thr Ala Cys Asp Val Gly
115         120         125

```

20130927_CL5646WOPCT_ST25. txt

Thr Tyr Val Ala Asp Met Asp Val Val Ala Arg Glu Cys Arg Trp Thr
130 135 140

Thr Gly Arg Ser Pro Glu Asn Gly Gly Ala Gly Asp Ser Ser Val Leu
145 150 155 160

Thr Ser Phe Gly Val Tyr Gln Gly Met Arg Ala Ala Ala Gln His Leu
165 170 175

Trp Gly Asp Pro Thr Leu Arg Asp Arg Thr Val Gly Ile Ala Gly Val
180 185 190

Gly Lys Val Gly His His Leu Val Glu His Leu Leu Ala Glu Gly Ala
195 200 205

His Val Val Val Thr Asp Val Arg Lys Asp Val Val Arg Gly Ile Thr
210 215 220

Glu Arg His Pro Ser Val Val Ala Val Ala Asp Thr Asp Ala Leu Ile
225 230 235 240

Arg Val Glu Asn Leu Asp Ile Tyr Ala Pro Cys Ala Leu Gly Gly Ala
245 250 255

Leu Asn Asp Asp Thr Val Pro Val Leu Thr Ala Lys Val Val Cys Gly
260 265 270

Ala Ala Asn Asn Gln Leu Ala His Pro Gly Val Glu Lys Asp Leu Ala
275 280 285

Asp Arg Gly Ile Leu Tyr Ala Pro Asp Tyr Val Val Asn Ala Gly Gly
290 295 300

Val Ile Gln Val Ala Asp Glu Leu His Gly Phe Asp Phe Asp Arg Cys
305 310 315 320

Lys Ala Lys Ala Ser Lys Ile Tyr Asp Thr Thr Leu Ala Ile Phe Ala
325 330 335

Arg Ala Lys Glu Asp Gly Ile Pro Pro Ala Ala Ala Ala Asp Arg Ile
340 345 350

Ala Glu Gln Arg Met Ala Glu Ala Arg Pro Arg Pro
355 360

<210> 82

<211> 1095

<212> DNA

<213> Streptomyces coelicolor

<400> 82

tcacggccgg ggacgggcct ccgcatccg ctgctcggcg atccggtcgg ccgccgcggc

60

cggcggaata ccgtcctcct tcgcacgtgc gaatatggcc agcgtggtgt cgtagatctt 120
 cgaggccttc gccttgacc ggtcgaagtc gaaccgtgc agtcgtcgg cgacctggat 180
 gacaccgccg gcgttcacca catagtccgg cgcgtagagg atcccgcggt cggcgaggtc 240
 ctctcgcacg cccgggtggg cgagctggtt gttggccgcg ccgcacacca ccttggcggt 300
 cagcaccggc acggtgtcgt cgttcagcgc gccgccgagc gcgcagggcg cgtagatgtc 360
 caggttctcc acccgatca gcgcgtcggg gtcggcgacg gcgaccaccg acgggtgccg 420
 ctccgtgatc ccgcgcacca cgtccttgcg cacgtccgtg acgacgacgt gggcgccctc 480
 ggcgagcagg tgctcgacca ggtggtggcc gaccttgccg acgcccgcga tgccgacggt 540
 gcggtcgcgc agcgtcgggt cgcacacag gtgctgggag gcggcccgca tgccctggta 600
 gacgccgaag gaggtgagca cggaggagtc gcccgcgccg ccgttctccg gggaacgccc 660
 ggtcgtccag cggcactcgc gggccacgac gtccatgtcg gcgacgtagg tgccgacgtc 720
 gcacgcggtg acgtagcggc cgcacagcga ggcgacgaac cggccgtagg cgaggagcag 780
 ctctcgcctc ttgatctgct ccggtcgcg gatgatcacg gccttgccgc caccgtggtc 840
 cagaccggcc atggcgttct tgtacgacat cccgcgggag aggttcagcg cgtcggcgac 900
 ggcctccgcc tcgctcgcgt acgggtagaa gcgggtaccg ccgagcgccg ggcccagggc 960
 ggtggagtgg agggcgatca cggccttgag gccgctggca cggtcctggc agagcacgac 1020
 ttgctcatgt ccccccctgat ccgagtggaa cagggtgtgc agtacatcag cagggtgcgc 1080
 gtttacgtcg gtcac 1095

<210> 83
 <211> 364
 <212> PRT
 <213> *Bacillus subtilis*

<400> 83

Met Glu Leu Phe Lys Tyr Met Glu Lys Tyr Asp Tyr Glu Gln Leu Val
1 5 10 15

Phe Cys Gln Asp Glu Gln Ser Gly Leu Lys Ala Ile Ile Ala Ile His
20 25 30

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

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

Met Thr Tyr Lys Asn Ala Ala Ala Gly Leu Asn Leu Gly Gly Gly Lys
65 70 75 80

Thr Val Ile Ile Gly Asp Pro Arg Lys Asp Lys Asn Glu Glu Met Phe
85 90 95

20130927_CL5646W0PCT_ST25. txt

Arg Ala Phe Gly Arg Tyr Ile Gln Gly Leu Asn Gly Arg Tyr Ile Thr
100 105 110

Ala Glu Asp Val Gly Thr Thr Val Glu Asp Met Asp Ile Ile His Asp
115 120 125

Glu Thr Asp Tyr Val Thr Gly Ile Ser Pro Ala Phe Gly Ser Ser Gly
130 135 140

Asn Pro Ser Pro Val Thr Ala Tyr Gly Val Tyr Arg Gly Met Lys Ala
145 150 155 160

Ala Ala Lys Ala Ala Phe Gly Thr Asp Ser Leu Glu Gly Lys Thr Ile
165 170 175

Ala Val Gln Gly Val Gly Asn Val Ala Tyr Asn Leu Cys Arg His Leu
180 185 190

His Glu Glu Gly Ala Asn Leu Ile Val Thr Asp Ile Asn Lys Gln Ser
195 200 205

Val Gln Arg Ala Val Glu Asp Phe Gly Ala Arg Ala Val Asp Pro Asp
210 215 220

Asp Ile Tyr Ser Gln Asp Cys Asp Ile Tyr Ala Pro Cys Ala Leu Gly
225 230 235 240

Ala Thr Ile Asn Asp Asp Thr Ile Lys Gln Leu Lys Ala Lys Val Ile
245 250 255

Ala Gly Ala Ala Asn Asn Gln Leu Lys Glu Thr Arg His Gly Asp Gln
260 265 270

Ile His Glu Met Gly Ile Val Tyr Ala Pro Asp Tyr Val Ile Asn Ala
275 280 285

Gly Gly Val Ile Asn Val Ala Asp Glu Leu Tyr Gly Tyr Asn Ala Glu
290 295 300

Arg Ala Leu Lys Lys Val Glu Gly Ile Tyr Gly Asn Ile Glu Arg Val
305 310 315 320

Leu Glu Ile Ser Gln Arg Asp Gly Ile Pro Ala Tyr Leu Ala Ala Asp
325 330 335

Arg Leu Ala Glu Glu Arg Ile Glu Arg Met Arg Arg Ser Arg Ser Gln
340 345 350

Phe Leu Gln Asn Gly His Ser Val Leu Ser Arg Arg
355 360

<210> 84
 <211> 364
 <212> PRT
 <213> Bacillus subtilis

<400> 84

Met Glu Leu Phe Lys Tyr Met Glu Lys Tyr Asp Tyr Glu Gln Leu Val
 1 5 10 15

Phe Cys Gln Asp Glu Gln Ser Gly Leu Lys Ala Ile Ile Ala Ile His
 20 25 30

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

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

Met Thr Tyr Lys Asn Ala Ala Ala Gly Leu Asn Leu Gly Gly Gly Lys
 65 70 75 80

Thr Val Ile Ile Gly Asp Pro Arg Lys Asp Lys Asn Glu Glu Met Phe
 85 90 95

Arg Ala Phe Gly Arg Tyr Ile Gln Gly Leu Asn Gly Arg Tyr Ile Thr
 100 105 110

Ala Glu Asp Val Gly Thr Thr Val Glu Asp Met Asp Ile Ile His Asp
 115 120 125

Glu Thr Asp Tyr Val Thr Gly Ile Ser Pro Ala Phe Gly Ser Ser Gly
 130 135 140

Asn Pro Ser Pro Val Thr Ala Tyr Gly Val Tyr Arg Gly Met Lys Ala
 145 150 155 160

Ala Ala Lys Ala Ala Phe Gly Thr Asp Ser Leu Glu Gly Lys Thr Ile
 165 170 175

Ala Val Gln Gly Val Gly Asn Val Ala Tyr Asn Leu Cys Arg His Leu
 180 185 190

His Glu Glu Gly Ala Asn Leu Ile Val Thr Asp Ile Asn Lys Gln Ser
 195 200 205

Val Gln Arg Ala Val Glu Asp Phe Gly Ala Arg Ala Val Asp Pro Asp
 210 215 220

Asp Ile Tyr Ser Gln Asp Cys Asp Ile Tyr Ala Pro Cys Ala Leu Gly
 225 230 235 240

Ala Thr Ile Asn Asp 245 Thr Ile Lys Gln Leu Lys Ala Lys Val 255 Ile

Ala Gly Ala Ala 260 Asn Asn Gln Leu Lys 265 Glu Thr Arg His Gly 270 Asp Gln

Ile His Glu 275 Met Gly Ile Val Tyr 280 Ala Pro Asp Tyr Val 285 Ile Asn Ala

Gly Gly 290 Val Ile Asn Val Ala 295 Asp Glu Leu Tyr Gly 300 Tyr Asn Ala Glu

Arg Ala Leu Lys Lys Val 310 Glu Gly Ile Tyr Gly 315 Asn Ile Glu Arg Val 320

Leu Glu Ile Ser Gln 325 Arg Asp Gly Ile Pro 330 Ala Tyr Leu Ala Ala 335 Asp

Arg Leu Ala Glu 340 Glu Arg Ile Glu Arg 345 Met Arg Arg Ser Arg 350 Ser Gln

Phe Leu Gln 355 Asn Gly His Ser Val 360 Leu Ser Arg Arg

<210> 85
<211> 594
<212> PRT
<213> Streptomyces viridifaciens

<400> 85

Met Ser Thr Ser Ser 5 Ala Ser Ser Gly Pro 10 Asp Leu Pro Phe Gly Pro 15

Glu Asp Thr Pro 20 Trp Gln Lys Ala Phe 25 Ser Arg Leu Arg Ala Val Asp 30

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

Met Asp Ile Pro Glu Ile 55 Pro Phe Ser Lys Val 60 Gln Ile Pro Pro Asp

Gly Met Asp Glu Gln 70 Gln Tyr Ala Glu Ala Glu 75 Ser Leu Phe Arg Arg 80

Tyr Val Asp Ala 85 Gln Thr Arg Asn Phe Ala 90 Gly Tyr Gln Val Thr Ser 95

Asp Leu Asp Tyr 100 Gln His Leu Ser His 105 Tyr Leu Asn Arg His 110 Leu Asn

Asn Val Gly Asp Pro Tyr Glu Ser Ser Ser Tyr Thr Leu Asn Ser Lys

115

120

125

Val Leu Glu Arg Ala Val Leu Asp Tyr Phe Ala Ser Leu Trp Asn Ala
 130 135 140
 Lys Trp Pro His Asp Ala Ser Asp Pro Glu Thr Tyr Trp Gly Tyr Val
 145 150 155 160
 Leu Thr Met Gly Ser Ser Glu Gly Asn Leu Tyr Gly Leu Trp Asn Ala
 165 170 175
 Arg Asp Tyr Leu Ser Gly Lys Leu Leu Arg Arg Glu His Arg Glu Ala
 180 185 190
 Gly Gly Asp Lys Ala Ser Val Val Tyr Thr Glu Ala Leu Arg His Glu
 195 200 205
 Gly Glu Ser Pro His Ala Tyr Glu Pro Val Ala Phe Phe Ser Glu Asp
 210 215 220
 Thr His Tyr Ser Leu Thr Lys Ala Val Arg Val Leu Gly Ile Asp Thr
 225 230 235 240
 Phe His Ser Ile Gly Ser Ser Arg Tyr Pro Asp Glu Asn Pro Leu Gly
 245 250 255
 Pro Gly Thr Pro Trp Pro Thr Glu Val Pro Ser Val Asp Gly Ala Ile
 260 265 270
 Asp Val Asp Lys Leu Ala Ser Leu Val Arg Phe Phe Ala Ser Lys Gly
 275 280 285
 Tyr Pro Ile Leu Val Ser Leu Asn Tyr Gly Ser Thr Phe Lys Gly Ala
 290 295 300
 Tyr Asp Asp Val Pro Ala Val Ala Glu Ala Val Arg Asp Ile Cys Thr
 305 310 315 320
 Glu Tyr Gly Leu Asp Arg Arg Arg Val Tyr His Asp Arg Ser Lys Asp
 325 330 335
 Ser Asp Phe Asp Glu Arg Ser Gly Phe Trp Ile His Ile Asp Ala Ala
 340 345 350
 Leu Gly Ala Gly Tyr Ala Pro Tyr Leu Glu Met Ala Arg Asp Ala Gly
 355 360 365
 Met Val Glu Glu Ala Pro Pro Val Phe Asp Phe Arg Leu Pro Glu Val
 370 375 380
 His Ser Leu Thr Met Ser Gly His Lys Trp Met Gly Thr Pro Trp Ala
 Page 128

385 390 395 400

Cys Gly Val Tyr Met Thr Arg Thr Gly Leu Gln Met Thr Pro Pro Lys
405 410 415

Ser Ser Glu Tyr Ile Gly Ala Ala Asp Thr Thr Phe Ala Gly Ser Arg
420 425 430

Asn Gly Phe Ser Ser Leu Leu Leu Trp Asp Tyr Leu Ser Arg His Ser
435 440 445

Tyr Asp Asp Leu Val Arg Leu Ala Ala Asp Cys Asp Arg Leu Ala Gly
450 455 460

Tyr Ala His Asp Arg Leu Leu Thr Leu Gln Asp Lys Leu Gly Met Asp
465 470 475 480

Leu Trp Val Ala Arg Ser Pro Gln Ser Leu Thr Val Arg Phe Arg Gln
485 490 495

Pro Cys Ala Asp Ile Val Arg Lys Tyr Ser Leu Ser Cys Glu Thr Val
500 505 510

Tyr Glu Asp Asn Glu Gln Arg Thr Tyr Val His Leu Tyr Ala Val Pro
515 520 525

His Leu Thr Arg Glu Leu Val Asp Glu Leu Val Arg Asp Leu Arg Gln
530 535 540

Pro Gly Ala Phe Thr Asn Ala Gly Ala Leu Glu Gly Glu Ala Trp Ala
545 550 555 560

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

Gly Ala Leu Ser Ala Pro Ala Ser Gly Pro Arg Ser Glu Asp Gly Gly
580 585 590

Gly Ser

<210> 86
<211> 1785
<212> DNA
<213> Streptomyces vi ri di faci ens

<400> 86
gtgtcaactt cctccgcttc ttccggggccg gacctcccct tcgggcccga ggacacgcca 60
tggcagaagg ccttcagcag gctgcgggcg gtggatggcg tgccgcgcgt caccgcgccg 120
tccagtgatc cgcgtgaggt ctacatggac atcccggaga tccccttctc caaggtccag 180
atcccccccg acggaatgga cgagcagcag tacgcagagg ccgagagcct cttccgccgc 240

20130927_CL5646WOPCT_ST25.txt

tacgtagacg cccagacccg caacttcgcg ggataccagg tcaccagcga cctcgactac 300
cagcacctca gtcactatct caaccggcat ctgaacaacg tcggcgatcc ctatgagtc 360
agctcctaca cgctgaactc caaggtcctt gagcgagccg ttctcgacta cttcgccctc 420
ctgtggaacg ccaagtggcc ccatgacgca agcgatccgg aaacgtactg gggttacgtg 480
ctgaccatgg gtcacagcga aggcaacctg tacgggttgt ggaacgcacg ggactatctg 540
tcgggcaagc tgctgcggcg ccagcaccgg gaggccggcg gcgacaaggc ctcggtcgtc 600
tacacgcaag cgctgcgaca cgaagggcag agtccgcatg cctacgagcc ggtggcgttc 660
ttctcgagg acacgacta ctcgctcagc aaggccgtgc gggttctggg catcgacacc 720
ttccacagca tcggcagcag tcggtatccg gacgagaacc cgctgggccc cggcactccg 780
tgcccgaccg aagtgccctc ggttgacggg gccatcgatg tcgacaaact cgcctcgttg 840
gtccgcttct tcgccagcaa gggctacccg atactggtca gcctcaacta cgggtcaacg 900
ttcaagggcg cctacgacga cgtcccggcc gtggcacagg ccgtgcggga catctgcacg 960
gaatacggtc tggatcggcg gcgggtatac cagcaccgca gtaaggacag tgacttcgac 1020
gagcgacgcg gcttctggat ccacatcgat gccgccctgg gggcgggcta cgctccctac 1080
ctgcagatgg cccgggatgc cggcatggtc gaggaggcgc cgcccgtttt cgacttccgg 1140
ctcccggagg tgactcgct gaccatgagc ggccacaagt ggatgggaac accgtgggca 1200
tgcggtgtct acatgacacg gaccgggctg cagatgaccc cgccgaagtc gtccgagtac 1260
atcggggcgg ccgacaccac cttcggggc tcccgcaacg gcttctcgtc actgctgctg 1320
tgggactacc tgtcccggca ttcgtatgac gatctggtgc gcctggccgc cgactgcgac 1380
cggctggccg gctacgccc cgaccgggtg ctgaccttg aggacaaact cggcatggat 1440
ctgtgggtcg cccgcagccc gcagtccctc acggtgcgct tccgtcagcc atgtgcagac 1500
atcgtccgca agtactcgct gtcgtgtgag acggtctacg aagacaacga gcaacggacc 1560
tacgtacatc tctacgccgt tccccacctc actcggaac tcgtggatga gctcgtgcgc 1620
gatctgcgcc agcccggagc cttaccaac gctggtgcac tggaggggga ggcctgggcc 1680
ggggtgatcg atgccctcg ccgcccggac cccgacggaa cctatgccgg cgccttgagc 1740
gctccggctt ccggcccccg ctccgaggac ggcgccggga gctga 1785

<210> 87

<211> 440

<212> PRT

<213> Al cal i genes deni tri fi cans

<400> 87

Met Ser Ala Ala Lys Leu Pro Asp Leu Ser Hi s Leu Trp Met Pro Phe
1 5 10 15

Thr Ala Asn Arg Gln Phe Lys Ala Asn Pro Arg Leu Leu Ala Ser Ala
20 25 30

20130927_CL5646WOPCT_ST25. txt

Lys Gly Met Tyr Tyr Thr Ser Phe Asp Gly Arg Gl n I l e Leu Asp Gly
 35 40 45
 Thr Al a Gly Leu Trp Cys Val Asn Al a Gly Hi s Cys Arg Gl u Gl u I l e
 50 55 60
 Val Ser Al a I l e Al a Ser Gl n Al a Gly Val Met Asp Tyr Al a Pro Gly
 65 70 75 80
 Phe Gl n Leu Gly Hi s Pro Leu Al a Phe Gl u Al a Al a Thr Al a Val Al a
 85 90 95
 Gly Leu Met Pro Gl n Gly Leu Asp Arg Val Phe Phe Thr Asn Ser Gly
 100 105 110
 Ser Gl u Ser Val Asp Thr Al a Leu Lys I l e Al a Leu Al a Tyr Hi s Arg
 115 120 125
 Al a Arg Gly Gl u Al a Gl n Arg Thr Arg Leu I l e Gly Arg Gl u Arg Gly
 130 135 140
 Tyr Hi s Gly Val Gly Phe Gly Gly I l e Ser Val Gly Gly I l e Ser Pro
 145 150 155 160
 Asn Arg Lys Thr Phe Ser Gly Al a Leu Leu Pro Al a Val Asp Hi s Leu
 165 170 175
 Pro Hi s Thr Hi s Ser Leu Gl u Hi s Asn Al a Phe Thr Arg Gly Gl n Pro
 180 185 190
 Gl u Trp Gly Al a Hi s Leu Al a Asp Gl u Leu Gl u Arg I l e I l e Al a Leu
 195 200 205
 Hi s Asp Al a Ser Thr I l e Al a Al a Val I l e Val Gl u Pro Met Al a Gly
 210 215 220
 Ser Thr Gly Val Leu Val Pro Pro Lys Gly Tyr Leu Gl u Lys Leu Arg
 225 230 235 240
 Gl u I l e Thr Al a Arg Hi s Gly I l e Leu Leu I l e Phe Asp Gl u Val I l e
 245 250 255
 Thr Al a Tyr Gly Arg Leu Gly Gl u Al a Thr Al a Al a Al a Tyr Phe Gly
 260 265 270
 Val Thr Pro Asp Leu I l e Thr Met Al a Lys Gly Val Ser Asn Al a Al a
 275 280 285
 Val Pro Al a Gly Al a Val Al a Val Arg Arg Gl u Val Hi s Asp Al a I l e
 290 295 300

20130927_CL5646WOPCT_ST25. txt

Val Asn Gly Pro Gln Gly Gly Ile Glu Phe Phe His Gly Tyr Thr Tyr
305 310 315 320

Ser Ala His Pro Leu Ala Ala Ala Ala Val Leu Ala Thr Leu Asp Ile
325 330 335

Tyr Arg Arg Glu Asp Leu Phe Ala Arg Ala Arg Lys Leu Ser Ala Ala
340 345 350

Phe Glu Glu Ala Ala His Ser Leu Lys Gly Ala Pro His Val Ile Asp
355 360 365

Val Arg Asn Ile Gly Leu Val Ala Gly Ile Glu Leu Ser Pro Arg Glu
370 375 380

Gly Ala Pro Gly Ala Arg Ala Ala Glu Ala Phe Gln Lys Cys Phe Asp
385 390 395 400

Thr Gly Leu Met Val Arg Tyr Thr Gly Asp Ile Leu Ala Val Ser Pro
405 410 415

Pro Leu Ile Val Asp Glu Asn Gln Ile Gly Gln Ile Phe Glu Gly Ile
420 425 430

Gly Lys Val Leu Lys Glu Val Ala
435 440

<210> 88
<211> 1947
<212> DNA
<213> Al cal i genes deni tri fi cans

<400> 88
ttcgatggcg cgctgcacgg cggccaccag ctgctccacc aggggtgggc gcctgcccgc 60
gcgcgcggtc gggctggaaa tcgatcatgg atgaatctat acagttgtca tgattgcaac 120
tatacagtta gcccgttttg cggcaattgt atattttcat tcgctcgtgg acgtccgaga 180
atcggtttga tcgcgccgcc cgccccttc cgcgacgcgg cgtttctttt cctccggagt 240
ctcccatga gcgctgcaa actgcccagc ctgtcccacc tctggatgcc cttaccgcc 300
aaccggcagt tcaaggcgaa ccccgccctg ctggcctcgg ccaagggcat gtactacacg 360
tctttcgacg gccgccagat cctggacggc acggccggcc tgtggtgcgt gaacgccggc 420
cactgccgcg aagaaatcgt ctccgccatc gccagccagg ccggcgatcat ggactacgcg 480
ccggggttcc agctcggcca cccgctggcc ttcgaggccg ccaccgccgt ggccggcctg 540
atgccgcagg gcctggaccg cgtgttcttc accaattcgg gctccgaatc ggtggacacc 600
gcgctgaaga tcgccctggc ctaccaccgc gcgcgcggcg aggcgcagcg caccgcctc 660
atcgggcgcg agcgcggcta ccacggcgtg ggcttcggcg gcatttccgt gggcggcatc 720
tcgccaacc gcaagacctt ctccggcgcg ctgctgccgg ccgtggacca cctgccgcac 780

20130927_CL5646WOPCT_ST25. txt

```

accacagcc tggaaacaaa cgccttcacg cgcggccagc ccgagtgggg cgcgcacctg      840
gccgacgagt tggaaacgat catcgccctg cagcagccct ccaccatcgc ggccgtgata      900
gtcgagccca tggccggctc caccggcgtg ctggtcccg ccaagggcta tctcgaaaaa      960
ctgcgcgaaa tcaccgcccg ccacggcatt ctgctgatct tcgacgaagt catcaccgcg     1020
tacggccgcc tgggcgaggc caccgccgcg gcctatttcg gcgtaacgcc cgacctatc      1080
accatggcca agggcgtag caacgccgcc gttccggccg gcgccgtcgc ggtgcgccgc      1140
gaagtgcata acgcatcgt caacggaccg caaggcgga tcgagttctt ccacggctac      1200
acctactcgg cccaccgct ggccgccgcc gccgtgctcg ccacgtgga catctaccgc      1260
cgcgaagacc tgttcgcccg cggccgcaag ctgtcggccg cgttcgagga agccgccac      1320
agcctcaagg gcgcgccga cgtcatcgac gtgcgaaca tcggcctggt ggccggcatc      1380
gagctgtcgc cgcgcaagg cggccgggc gcgcgcgcc cgaagcctt ccagaaatgc      1440
ttcgacaccg gcctcatggt gcgtacacg ggcgacatcc tcgcggtgtc gcctccgctc      1500
atcgtcgacg aaaaccagat cggccagatc ttcgaggga tcggcaaggt gctcaaggaa      1560
gtggcttagg gtgaacacgc cctgagccgg ccccgccagg aaacgcgcc cgcgcgggcg      1620
gcgcgtccat cgaactcccg catcgagctt ttgcattcat gaagaaaatc acgcatttca      1680
tcaacggcca gcccacgaa ggccgcagca accgctacac cgagggttc aaccggcca      1740
cgggcgagtc gtctcctga tctgcctggg cggggccgaa gaagtggacc tggccgtggc      1800
ggccgccgc gcggccttc ccgctggtc cgaaacgcc gcgtcaagc gcgcgcgct      1860
gctgttaac ttcaaggcg tgctggaaa gcaccaggac gagctggccg cgctcatcac      1920
gcgcgagcac ggcaaggtgt tttccga      1947

```

<210> 89
 <211> 443
 <212> PRT
 <213> *Ralstonia eutropha*

<400> 89

```

Met Asp Ala Ala Lys Thr Val Ile Pro Asp Leu Asp Ala Leu Trp Met
1           5           10           15

```

```

Pro Phe Thr Ala Asn Arg Gln Tyr Lys Ala Ala Pro Arg Leu Leu Ala
          20           25           30

```

```

Ser Ala Ser Gly Met Tyr Tyr Thr Thr His Asp Gly Arg Gln Ile Leu
35           40           45

```

```

Asp Gly Cys Ala Gly Leu Trp Cys Val Ala Ala Gly His Cys Arg Lys
50           55           60

```

```

Glu Ile Ala Glu Ala Val Ala Arg Gln Ala Ala Thr Leu Asp Tyr Ala
65           70           75           80

```

20130927_CL5646WOPCT_ST25. txt

Pro Pro Phe Gln Met Gly His Pro Leu Ser Phe Glu Ala Ala Thr Lys
85 90 95

Val Ala Ala Ile Met Pro Gln Gly Leu Asp Arg Ile Phe Phe Thr Asn
100 105 110

Ser Gly Ser Glu Ser Val Asp Thr Ala Leu Lys Ile Ala Leu Ala Tyr
115 120 125

His Arg Ala Arg Gly Glu Gly Gln Arg Thr Arg Phe Ile Gly Arg Glu
130 135 140

Arg Gly Tyr His Gly Val Gly Phe Gly Gly Met Ala Val Gly Gly Ile
145 150 155 160

Gly Pro Asn Arg Lys Ala Phe Ser Ala Asn Leu Met Pro Gly Thr Asp
165 170 175

His Leu Pro Ala Thr Leu Asn Ile Ala Glu Ala Ala Phe Ser Lys Gly
180 185 190

Gln Pro Thr Trp Gly Ala His Leu Ala Asp Glu Leu Glu Arg Ile Val
195 200 205

Ala Leu His Asp Pro Ser Thr Ile Ala Ala Val Ile Val Glu Pro Leu
210 215 220

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

Leu Arg Glu Ile Thr Thr Lys His Gly Ile Leu Leu Ile Phe Asp Glu
245 250 255

Val Ile Thr Ala Phe Gly Arg Leu Gly Thr Ala Thr Ala Ala Glu Arg
260 265 270

Phe Lys Val Thr Pro Asp Leu Ile Thr Met Ala Lys Ala Ile Asn Asn
275 280 285

Ala Ala Val Pro Met Gly Ala Val Ala Val Arg Arg Glu Val His Asp
290 295 300

Thr Val Val Asn Ser Ala Ala Pro Gly Ala Ile Glu Leu Ala His Gly
305 310 315 320

Tyr Thr Tyr Ser Gly His Pro Leu Ala Ala Ala Ala Ala Ile Ala Thr
325 330 335

Leu Asp Leu Tyr Gln Arg Glu Asn Leu Phe Gly Arg Ala Ala Glu Leu
340 345 350

20130927_CL5646WOPCT_ST25. txt

Ser Pro Val Phe Glu Ala Ala Val His Ser Val Arg Ser Ala Pro His
355 360 365

Val Lys Asp Ile Arg Asn Leu Gly Met Val Ala Gly Ile Glu Leu Glu
370 375 380

Pro Arg Pro Gly Gln Pro Gly Ala Arg Ala Tyr Glu Ala Phe Leu Lys
385 390 395 400

Cys Leu Glu Arg Gly Val Leu Val Arg Tyr Thr Gly Asp Ile Leu Ala
405 410 415

Phe Ser Pro Pro Leu Ile Ile Ser Glu Ala Gln Ile Ala Glu Leu Phe
420 425 430

Asp Thr Val Lys Gln Ala Leu Gln Glu Val Gln
435 440

<210> 90
<211> 1341
<212> DNA
<213> Ralstonia eutropha

<400> 90
atggccgact cacccaacaa cctcgctcac gaacatcctt cacttgaaca ctattggatg 60
ccttttaccg ccaatcgcca attcaaagcg agccctcggt tactcgcca agctgaaggt 120
atgtattaca cagatatcaa tggcaacaag gtattagact ctacagcggg cttatggtgt 180
tgtaatgctg gccatggtcg ccgtgagatc agtgaagccg tcagcaaaca aattcggcag 240
atggattacg ctccctcctt ccaaattgggc catcccatcg cttttgaact ggccgaacgt 300
ttaaccgaac tcagcccaga aggactcaac aaagtattct ttaccaactc aggctctgag 360
tcggttgata ccgcgctaaa aatggctctt tgctaccata gagccaatgg ccaagcgtca 420
cgcacccgct ttattggccg tgaaatgggt taccatggcg taggatttg tgggatctcg 480
gtgggtggtt taagcaataa ccgtaaagcc ttcagcggcc agctattgca aggcgtggat 540
cacctgcccc acaccttaga cattcaacat gccgccttta gtcgtggctt accgagcctc 600
ggtgctgaaa aagctgaggt attagaacaa ttagtcacac tccatggcgc cgaaaatatt 660
gccgccgtta ttgttgaacc catgtcaggt tctgcagggg taattttacc acctcaaggc 720
tacttaaaac gcttacgtga aatcactaaa aaacacggca tcttattgat tttcgatgaa 780
gtcattaccg catttggccg ttaggtgca gcattcgcca gccaacgttg gggcgttatt 840
ccagacataa tcaccacggc taaagccatt aataatggcg ccatcccat gggcgcagtg 900
ttgttacagg attatatcca cgatacttgc atgcaagggc caaccgaact gattgaattt 960
ttccacggtt atacctattc gggccacca gtcgccgcag cagcagcact cgccacgctc 1020
tccatctacc aaaacgagca actgtttgag cgcagttttg agcttgagcg gtatttcgaa 1080
gaagccgttc atagcctcaa agggttaccg aatgtgattg atattcgcaa caccggatta 1140

gtcgcgggtt tccagctagc accgaatagc caaggtgttg gtaaacgcgg atacagcgtg 1200
 ttcgagcatt gtttccatca aggcacactc gtgcgggcaa cgggcgatat tatcgccatg 1260
 tccccaccac tcattgttga gaaacatcag attgaccaaa tggtaaatag ccttagcgat 1320
 gcaattcacg ccgttgatg a 1341

<210> 91
 <211> 446
 <212> PRT
 <213> Shewanella oneidensis
 <400> 91

Met Ala Asp Ser Pro Asn Asn Leu Ala His Glu His Pro Ser Leu Glu
 1 5 10 15

His Tyr Trp Met Pro Phe Thr Ala Asn Arg Gln Phe Lys Ala Ser Pro
 20 25 30

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

Asn Lys Val Leu Asp Ser Thr Ala Gly Leu Trp Cys Cys Asn Ala Gly
 50 55 60

His Gly Arg Arg Glu Ile Ser Glu Ala Val Ser Lys Gln Ile Arg Gln
 65 70 75 80

Met Asp Tyr Ala Pro Ser Phe Gln Met Gly His Pro Ile Ala Phe Glu
 85 90 95

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

Phe Phe Thr Asn Ser Gly Ser Glu Ser Val Asp Thr Ala Leu Lys Met
 115 120 125

Ala Leu Cys Tyr His Arg Ala Asn Gly Gln Ala Ser Arg Thr Arg Phe
 130 135 140

Ile Gly Arg Glu Met Gly Tyr His Gly Val Gly Phe Gly Gly Ile Ser
 145 150 155 160

Val Gly Gly Leu Ser Asn Asn Arg Lys Ala Phe Ser Gly Gln Leu Leu
 165 170 175

Gln Gly Val Asp His Leu Pro His Thr Leu Asp Ile Gln His Ala Ala
 180 185 190

Phe Ser Arg Gly Leu Pro Ser Leu Gly Ala Glu Lys Ala Glu Val Leu
 195 200 205

Glu Gln Leu Val Thr Leu His Gly Ala Glu Asn Ile Ala Ala Val Ile
210 215 220

Val Glu Pro Met Ser Gly Ser Ala Gly Val Ile Leu Pro Pro Gln Gly
225 230 235 240

Tyr Leu Lys Arg Leu Arg Glu Ile Thr Lys Lys His Gly Ile Leu Leu
245 250 255

Ile Phe Asp Glu Val Ile Thr Ala Phe Gly Arg Val Gly Ala Ala Phe
260 265 270

Ala Ser Gln Arg Trp Gly Val Ile Pro Asp Ile Ile Thr Thr Ala Lys
275 280 285

Ala Ile Asn Asn Gly Ala Ile Pro Met Gly Ala Val Phe Val Gln Asp
290 295 300

Tyr Ile His Asp Thr Cys Met Gln Gly Pro Thr Glu Leu Ile Glu Phe
305 310 315 320

Phe His Gly Tyr Thr Tyr Ser Gly His Pro Val Ala Ala Ala Ala Ala
325 330 335

Leu Ala Thr Leu Ser Ile Tyr Gln Asn Glu Gln Leu Phe Glu Arg Ser
340 345 350

Phe Glu Leu Glu Arg Tyr Phe Glu Glu Ala Val His Ser Leu Lys Gly
355 360 365

Leu Pro Asn Val Ile Asp Ile Arg Asn Thr Gly Leu Val Ala Gly Phe
370 375 380

Gln Leu Ala Pro Asn Ser Gln Gly Val Gly Lys Arg Gly Tyr Ser Val
385 390 395 400

Phe Glu His Cys Phe His Gln Gly Thr Leu Val Arg Ala Thr Gly Asp
405 410 415

Ile Ile Ala Met Ser Pro Pro Leu Ile Val Glu Lys His Gln Ile Asp
420 425 430

Gln Met Val Asn Ser Leu Ser Asp Ala Ile His Ala Val Gly
435 440 445

<210> 92

<211> 1341

<212> DNA

<213> Shewanella oneidensis

<400> 92

atggccgact cacccaacaa cctcgctcac gaacatcctt cacttgaaca ctattggatg

60

```

ccttttaccg ccaatcgcca attcaaagcg agccctcggt tactcgccca agctgaaggt      120
atgtattaca cagatatcaa tggcaacaag gtattagact ctacagcggg cttatggtgt      180
tgtaatgctg gccatggtcg ccgtgagatc agtgaagccg tcagcaaaca aattcggcag      240
atggattacg ctccctcctt ccaaatgggc catcccatcg cttttgaact ggccgaacgt      300
ttaaccgaac tcagcccaga aggactcaac aaagtattct ttaccaactc aggctctgag      360
tcggttgata ccgcgctaaa aatggctctt tgctaccata gagccaatgg ccaagcgtca      420
cgcacccgct ttattggccg tgaaatgggt taccatggcg taggatttgg tgggatctcg      480
gtgggtgggt taagcaataa ccgtaaagcc ttacagcggc agctattgca aggcgtggat      540
cacctgcccc acaccttaga cattcaacat gccgccttta gtcgtggctt accgagcctc      600
ggtgctgaaa aagctgaggt attagaacaa ttagtcacac tccatggcgc cgaaaatatt      660
gccgccgcta ttgttggaac catgtcaggt tctgcagggg taattttacc acctcaaggc      720
tacttaaaac gcttacgtga aatcactaaa aaacacggca tcttattgat tttcgatgaa      780
gtcattaccg catttggccg ttaggtgca gcattcgcca gccaacgttg gggcgttatt      840
ccagacataa tcaccacggc taaagccatt aataatggcg ccatcccat gggcgcagtg      900
tttgtacagg attatatcca cgatacttgc atgcaagggc caaccgaact gattgaattt      960
ttccacgggt atacctattc gggccacca gtcgccgcag cagcagcact cgccacgctc     1020
tccatctacc aaaacgagca actgtttgag cgcagttttg agcttgagcg gtatttcgaa     1080
gaagccgttc atagcctcaa agggttaccg aatgtgattg atattcgcaa caccggatta     1140
gtcgcggggt tccagctagc accgaatagc caaggtgttg gtaaacgcgg atacagcgtg     1200
ttcgagcatt gtttccatca aggcacactc gtgcgggcaa cgggcgatat tatcgccatg     1260
tccccaccac tcattgttga gaaacatcag attgaccaaa tggtaaatag ccttagcgat     1320
gcaattcacg ccgttgatg a                                     1341

```

```

<210> 93
<211> 448
<212> PRT
<213> Pseudomonas putida

```

```
<400> 93
```

```
Met Asn Met Pro Glu Thr Gly Pro Ala Gly Ile Ala Ser Gl n Leu Lys
1          5          10          15
```

```
Leu Asp Ala His Trp Met Pro Tyr Thr Ala Asn Arg Asn Phe Gl n Arg
20          25          30
```

```
Asp Pro Arg Leu Ile Val Ala Ala Glu Gly Asn Tyr Leu Val Asp Asp
35          40          45
```

```
His Gly Arg Lys Ile Phe Asp Ala Leu Ser Gly Leu Trp Thr Cys Gly
50          55          60
```

20130927_CL5646W0PCT_ST25. txt

Ala Gly His Thr Arg Lys Glu Ile Ala Asp Ala Val Thr Arg Gln Leu
65 70 75 80

Ser Thr Leu Asp Tyr Ser Pro Ala Phe Gln Phe Gly His Pro Leu Ser
85 90 95

Phe Gln Leu Ala Glu Lys Ile Ala Glu Leu Val Pro Gly Asn Leu Asn
100 105 110

His Val Phe Tyr Thr Asn Ser Gly Ser Glu Cys Ala Asp Thr Ala Leu
115 120 125

Lys Met Val Arg Ala Tyr Trp Arg Leu Lys Gly Gln Ala Thr Lys Thr
130 135 140

Lys Ile Ile Gly Arg Ala Arg Gly Tyr His Gly Val Asn Ile Ala Gly
145 150 155 160

Thr Ser Leu Gly Gly Val Asn Gly Asn Arg Lys Met Phe Gly Gln Leu
165 170 175

Leu Asp Val Asp His Leu Pro His Thr Val Leu Pro Val Asn Ala Phe
180 185 190

Ser Lys Gly Leu Pro Glu Glu Gly Gly Ile Ala Leu Ala Asp Glu Met
195 200 205

Leu Lys Leu Ile Glu Leu His Asp Ala Ser Asn Ile Ala Ala Val Ile
210 215 220

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

Tyr Leu Lys Arg Leu Arg Glu Ile Cys Thr Gln His Asn Ile Leu Leu
245 250 255

Ile Phe Asp Glu Val Ile Thr Gly Phe Gly Arg Met Gly Ala Met Thr
260 265 270

Gly Ser Glu Ala Phe Gly Val Thr Pro Asp Leu Met Cys Ile Ala Lys
275 280 285

Gln Val Thr Asn Gly Ala Ile Pro Met Gly Ala Val Ile Ala Ser Ser
290 295 300

Glu Ile Tyr Gln Thr Phe Met Asn Gln Pro Thr Pro Glu Tyr Ala Val
305 310 315 320

Glu Phe Pro His Gly Tyr Thr Tyr Ser Ala His Pro Val Ala Cys Ala
325 330 335

20130927_CL5646WOPCT_ST25. txt

Ala Gly Leu Ala Ala Leu Asp Leu Leu Gln Lys Glu Asn Leu Val Gln
340 345 350

Ser Ala Ala Glu Leu Ala Pro His Phe Glu Lys Leu Leu His Gly Val
355 360 365

Lys Gly Thr Lys Asn Ile Val Asp Ile Arg Asn Tyr Gly Leu Ala Gly
370 375 380

Ala Ile Gln Ile Ala Ala Arg Asp Gly Asp Ala Ile Val Arg Pro Tyr
385 390 395 400

Glu Ala Ala Met Lys Leu Trp Lys Ala Gly Phe Tyr Val Arg Phe Gly
405 410 415

Gly Asp Thr Leu Gln Phe Gly Pro Thr Phe Asn Thr Lys Pro Gln Glu
420 425 430

Leu Asp Arg Leu Phe Asp Ala Val Gly Glu Thr Leu Asn Leu Ile Asp
435 440 445

<210> 94
<211> 930
<212> DNA
<213> Pseudomonas putida

<400> 94
atgaccacga agaaagctga ttacatttgg ttcaatgggg agatggttcg ctgggaagac 60
gcgaagggtgc atgtgatgtc gcacgcgctg cactatggca cttcggtttt tgaaggcatc 120
cgttgctacg actcgcaaa aggaccgggt gtattccgcc atcgtgagca tatgcagcgt 180
ctgcatgact ccgcaaaaat ctatcgcttc ccggtttcgc agagcattga tgagctgatg 240
gaagcttgtc gtgacgtgat ccgcaaaaac aatctcacca gcgcctatat ccgtccgctg 300
atcttcgtcg gtgatgttgg catgggagta aaccgccag cgggatactc aaccgacgtg 360
attatcgctg ctttcccgtg gggagcgtat ctgggcgcag aagcgctgga gcaggggatc 420
gatgcgatgg tttcctcctg gaaccgcgca gcaccaaaca ccatcccgac ggcggaaaaa 480
gccggtggta actacctctc ttccctgctg gtgggtagcg aagcgcgccg ccacggttat 540
caggaaggta tcgcgctgga tgtgaacggt tatatctctg aaggcgcagg cgaaaacctg 600
tttgaagtga aagatggtgt gctgttcacc ccaccgttca cctcctccgc gctgccgggt 660
attaccctg atgcatcat caaactggcg aaagagctgg gaattgaagt acgtgagcag 720
gtgctgtcgc gcgaatccct gtacctggcg gatgaagtgt ttatgtccgg tacggcggca 780
gaaatcacgc cagtgcgcag cgtagacggt attcaggttg gcgaaggccg ttgtggcccc 840
gttaccaaac gcattcagca agccttcttc ggcctcttca ctggcgaaac cgaagataaa 900
tggggctggt tagatcaagt taatcaataa 930

<210> 95
 <211> 566
 <212> PRT
 <213> Streptomyces cinnamoniensis

<400> 95

Met Asp Ala Asp Ala Ile Glu Glu Gly Arg Arg Arg Trp Gln Ala Arg
 1 5 10 15

Tyr Asp Lys Ala Arg Lys Arg Asp Ala Asp Phe Thr Thr Leu Ser Gly
 20 25 30

Asp Pro Val Asp Pro Val Tyr Gly Pro Arg Pro Gly Asp Thr Tyr Asp
 35 40 45

Gly Phe Glu Arg Ile Gly Trp Pro Gly Glu Tyr Pro Phe Thr Arg Gly
 50 55 60

Leu Tyr Ala Thr Gly Tyr Arg Gly Arg Thr Trp Thr Ile Arg Gln Phe
 65 70 75 80

Ala Gly Phe Gly Asn Ala Glu Gln Thr Asn Glu Arg Tyr Lys Met Ile
 85 90 95

Leu Ala Asn Gly Gly Gly Gly Leu Ser Val Ala Phe Asp Met Pro Thr
 100 105 110

Leu Met Gly Arg Asp Ser Asp Asp Pro Arg Ser Leu Gly Glu Val Gly
 115 120 125

His Cys Gly Val Ala Ile Asp Ser Ala Ala Asp Met Glu Val Leu Phe
 130 135 140

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

Pro Ala Val Pro Val Phe Cys Met Tyr Leu Val Ala Ala Glu Arg Gln
 165 170 175

Gly Val Asp Pro Ala Val Leu Asn Gly Thr Leu Gln Thr Asp Ile Phe
 180 185 190

Lys Glu Tyr Ile Ala Gln Lys Glu Trp Leu Phe Gln Pro Glu Pro His
 195 200 205

Leu Arg Leu Ile Gly Asp Leu Met Glu His Cys Ala Arg Asp Ile Pro
 210 215 220

Ala Tyr Lys Pro Leu Ser Val Ser Gly Tyr His Ile Arg Glu Ala Gly
 225 230 235 240

Ala Thr Ala Ala Gln Glu Leu Ala Tyr Thr Leu Ala Asp Gly Phe Gly
 Page 141

515

520

525

Gly Ser Asn Met Ile Ala Pro Met Leu Glu Ala Val Arg Ala Glu Ala
 530 535 540

Thr Leu Gly Glu Ile Cys Gly Val Leu Arg Asp Glu Trp Gly Val Tyr
 545 550 555 560

Val Glu Pro Pro Gly Phe
 565

<210> 96
 <211> 4362
 <212> DNA
 <213> Streptomyces cinnamonensis

<400> 96
 tgaggcgctg gatcgctcg gagagcagct ggtaacggtc cgcgtggtac tcggccgggg 60
 tgcagccgtc cacgatgtgc gggatcgctg cgggctcgag gatcaccagg gcgggggctg 120
 cgccgatcgc gtcggcgaac gtgtccaccc agctccggtg ggcctccgca ctggccgcgc 180
 cgcccgcgga gtgctgaccg cagtcgcggt gcgggatgtt gtacgcgacg agtacggcgg 240
 tgcggtcctc cttgaccgcg ccccgctcg ccttcgcgac gtcgggcgcc ggatcgtccc 300
 cggccggcca cacggccatg gccggttcgg agatgcgcct gagcgtctcg gcgtcctcgg 360
 cgcgggcctg ttctccac tgcctgacct ggcgcgcggc ggggctgtcg gggtcgacct 420
 agaaggtgcc ggcggggggc cggcgctcg cggtggcggg cttgcgcacg gccgcctcct 480
 ccttcgtgcc gtcggacccc gggcttgagg aggagcagcc tgccgggagc ccgagggcgg 540
 cgagggccgc gagtgccgtg aacgtgcgga gcagccggtg catccagccc cttgggcga 600
 tggtagacagt gacggtcagt cagcccggca atcgttacat aaaggactat tcaagctctt 660
 gtgccacacc gcctccggtg ccgagcgcga acccggcggg caccagagcc ccgccgcggc 720
 cgcgagaccg tacgtacgac cgaattgcga gacggggctg accaccatat gaccggcggg 780
 taaggtcgat gccgtgccga agccgtcag cctccccttc gatcccatcg cccgcgccga 840
 cgagctctgg aagcagcgt ggggatcggg cccggccatg ggcgcgatca cctcgatcat 900
 gcgggcgcac cagatcctgc tcgccgaggt cgacgcggtc gtcaagccgt acggactgac 960
 cttcgcgcgc tacgaggcgc tgggtgtcct caccttctcg caggccggcg agttgccgat 1020
 gtcgaagatc ggcgagcggc tcatggtgca cccgacctcg gtcacgaaca ccgtggaccg 1080
 cctggtgaag tccggcctgg tcgacaagcg cccgaacccc aacgacggcc gcggcacgct 1140
 cgctccatc acggagaagg gccgcgaggt cgtcgaggcg gccaccgcg agctgatggc 1200
 gatggacttc gggctcgggg tgtacgacgc ggaggagtgc ggggagatct tcgcgatgct 1260
 gcggcccctg cgggtggcgg cgcgcgattt cgaggagcag tagggcccgc ccggtgagaa 1320
 gtgggatcgg gtcgtcccgg tacgggcggg ggcggcgaag atcgctgaa aagggcggtt 1380
 acgctcgtag ccatgaaacg cagcgtgctg acccgctacc gggatgatggc ctacgtcacc 1440

20130927_CL5646WOPCT_ST25. txt

gccgtcatgc	tcctcatcct	gtg'gcctgc	atggtggcca	agtacggctt	cgacaagggc	1500
gagggctctga	ccctcgtcgt	gtcgcagggtg	cacggcgctgc	tctacatcat	ctacctgac	1560
ttcgccttcg	acctgggctc	caaggcgaag	tggccgttcg	gcaagctgct	ctgggtgctg	1620
gtctcgggca	cgatcccgac	cgccgccttc	ttcgtcgagc	gcaaggtcgc	ccgtgacgtc	1680
gagccgctga	tcgccgacgg	ctccccggtc	accgcgaagg	cgtaaccgcg	accgccacgg	1740
acagggtccgt	ggcggttggc	catcgacttt	tactaggacg	tcctagtaaa	ttcgatggta	1800
tggacgctga	cgcgatcgag	gaaggccgcc	gacgctggca	ggcccgttac	gacaaggccc	1860
gcaagcgcga	cgcggaactc	accacgctct	ccggggaccc	cgtcgacccc	gtctacggcc	1920
cccggccccg	ggacacgtac	gacgggttcg	agcggatcgg	ctggccgggg	gagtaccctt	1980
tcacccgcgg	gctctacgcc	accgggtacc	gcggccgcac	ctggaccatc	cgccagttcg	2040
ccggcttcgg	caacgccgag	cagacgaacg	agcgctacaa	gatgatcctg	gccaacggcg	2100
gcggcggcct	ctccgtcgcc	ttcgacatgc	cgaccctcat	gggccgcgac	tccgacgacc	2160
cgcgctcgct	cggcgaggtc	ggccactgcg	gtgtcgccat	cgactccgcc	gccgacatgg	2220
aggtcctctt	caaggacatc	ccgctcggcg	acgtcacgac	gtccatgacc	atcagcgggc	2280
ccgccgtgcc	cgtcttctgc	atgtacctcg	tcgcggccga	gcgccagggc	gtcgacccgg	2340
ccgtcctcaa	cggcacgctg	cagaccgaca	tcttcaagga	gtacatcgcc	cagaaggagt	2400
ggctcttcca	gcccagagccg	cacctgcgcc	tcatcggcga	cctgatggag	cactgcgcgc	2460
gcgacatccc	cgcgatacaag	ccgctctcgg	tctccggcta	ccacatccgc	gaggccgggg	2520
cgacggccgc	gcaggagctc	gcgtacaccc	tcgcggacgg	cttcgggtac	gtggaactgg	2580
gcctctcgcg	cggcctggac	gtggacgtct	tcgcgcccgg	cctctccttc	ttcttcgacg	2640
cgcacgtcga	cttcttcgag	gagatcgcg	agtccgcgc	cgcacgccgc	atctgggcgc	2700
gctggctccg	ggacgagtag	ggagcgaaga	ccgagaaggc	acagtggctg	cgcttccaca	2760
cgcagaccgc	gggggtctcg	ctcacggccc	agcagccgta	caacaacgtg	gtgcggacgg	2820
cggtaggagg	cctcgccgcg	gtgctcggcg	gcacgaactc	cctgcacacc	aacgctctcg	2880
acgagaccct	tgccctcccc	agcgagcagg	ccgcggagat	cgcgctgcgc	accagcagg	2940
tgtgatgga	ggagaccggc	gtcgccaacg	tcgcggaccc	gctgggcggc	tcctggtaca	3000
tcgagcagct	caccgaccgc	atcgaggccg	acgccgagaa	gatcttcgag	cagatcaggg	3060
agcgggggcg	gcgggcctgc	cccgcggggc	agcaccgat	cgggccgatc	acctccggca	3120
tcctgcgcgg	catcgaggac	ggctggttca	ccggcgagat	cgccgagtcc	gccttccagt	3180
accagcggtc	cctggagaag	ggcgacaagc	gggtcgtcgg	cgtcaactgc	ctcgaaggct	3240
ccgtcaccgg	cgacctggag	atcctgcgcg	tcagccacga	ggtcgagcgc	gagcagggtc	3300
gggagcttgc	ggggcgcaag	gggcggcggtg	acgatgcgcg	ggtgcggggc	tcgctcgacg	3360
cgatgctcgc	cgctgcgcgg	gacgggtcga	acatgattgc	ccccatgctg	gaggcggtgc	3420
gggccgaggc	gaccctcggg	gagatctcgc	gggtgcttcg	cgatgagtgg	ggggtctacg	3480

20130927_CL5646WOPCT_ST25. txt

tggagccgcc cgggttctga gggcgcgctc cttttgcctg cgggtctgct gtggctggtc 3540
gcgcagttcc ccgcacccct gaaagacccc ggcgctttcc ctctctggct cgcctcgctc 3600
ctgtctgcgg ggccgtgggg gctggtcgcg cagttccccg cggccctgcc cgcacctgcg 3660
ccccgccgcc tgcattgccg cccacccctg acgggggcg tgggggcca ccctgacggg 3720
tgccgtcggg gcgtgccggg gtcttttagg ggcgcgggga actgcgcgag caacccccac 3780
ccacccgcag gtgcacgcgg agcggcggac gccccgcaga cgggggcaaa acgggcggag 3840
tgccccgcc cgccgggcgg cgcgaattcg taggtttaag gggcaggggt cagggcagggc 3900
gccgagccgg tcaaccgcc ccgtcccagg agaccccg tgcctgaccg gccacgccc 3960
caccgccgcc atcgccatcg gagccgccac cgccaccgtc ctccggcgcg tgctggtcgg 4020
cggctccggc gaggtgagtg cgagccgcc gcccgagccc aaggtccagg acgacttcga 4080
ctccctcggc cccgaggtgc gcgccgcgaa gctctccgac gggcggacgg cccactactc 4140
ggacacgggc gacaaggacg gcaagccggc cctgttcacg ggcggcaccg gcacgagcgc 4200
ccgcgcctcc cacatgaccg acttcttcg ctcgacgcgc gaggacctgg gcctgcgcct 4260
catctccgtg gagcgcaacg gcttcggcga caccgcgttc gacgagaagc tgggcaccgc 4320
cgacttcgcg aaggacgccc tcgaagtcct cgaccggctc gg 4362

<210> 97
<211> 136
<212> PRT
<213> Streptomyces cinnamonensis
<400> 97

Met Gly Val Ala Ala Gly Pro Ile Arg Val Val Val Ala Lys Pro Gly
1 5 10 15

Leu Asp Gly His Asp Arg Gly Ala Lys Val Ile Ala Arg Ala Leu Arg
20 25 30

Asp Ala Gly Met Glu Val Ile Tyr Thr Gly Leu His Gln Thr Pro Glu
35 40 45

Gln Val Val Asp Thr Ala Ile Gln Glu Asp Ala Asp Ala Ile Gly Leu
50 55 60

Ser Ile Leu Ser Gly Ala His Asn Thr Leu Phe Ala Arg Val Leu Glu
65 70 75 80

Leu Leu Lys Glu Arg Asp Ala Glu Asp Ile Lys Val Phe Gly Gly Gly
85 90 95

Ile Ile Pro Glu Ala Asp Ile Ala Pro Leu Lys Glu Lys Gly Val Ala
100 105 110

Glu Ile Phe Thr Pro Gly Ala Thr Thr Thr Ser Ile Val Glu Trp Val
Page 145

Arg Gly Asn Val Arg Gl n Al a Val
130 135

<210> 98
<211> 1643
<212> DNA
<213> Streptomyces ci nnamonensi s

<400> 98
gtcgacctcc cgtttggcgc acggaaggga ggctctgtcc cccgtgtgcc ctagggggag 60
tcgtggtcga ggagtcggct gtgcgatggc gatcccgcc accgccctgc ggtgactccg 120
tgccccgtt gcatcgccga tgcgcggtgt caccacgccg tgcggctgcc ggcgcggtgg 180
cccggcgtct cgttgcggct cccctcgcgc ctggtccgga tgcggagcgt gaacccttg 240
gttacggacg ggcgcgacg gaacgtgtcc cacgtgtgat ttccccctcg ctctccaccg 300
cgaaactgcc gcttgcgcga tgctggggat aacgttcgtt cacttccccg gccggtgcgg 360
tgcggggtat ctgtgccggg acagactttg tcggtacgga tatcgttaca tggaggcagt 420
gatgggtgtg gcagccgggc cgatccgcgt ggtggtcgcc aagccggggc tcgacgggca 480
cgatcgccgg gccaaagtga tcgcgcgggc gttgcgtgac gcgggtatgg aggtcatcta 540
caccgggctg caccagacgc ccgagcaggt ggtggacacc gcgatccagg aggacgccga 600
cgcgatcggc ctctccatcc tctccggagc gcacaacacg ctgttcgcgc gcgtgttgga 660
gctcttgaag gagcgggacg cggaggacat caaggtgttt ggtggcggca tcatcccga 720
ggcggacatc gcgccgtga aggagaagg cgtcgcggag atcttcacgc ccggggccac 780
caccacgtcg atcgtggagt gggttcgggg gaacgtgcga caggccgtct gaggcattcc 840
ccgtcgcccc tctgccgtgg tcggcgatcat atcggcggac atcgtctcgg tggacgtcat 900
ggcggcgggg ggagttcgtc gcgtatcgcc gcgcggaggc gcagggtggt gaccaggcgc 960
tggaaacgtt ccgaccagta gctgcccgcg ccgggtgacg cgtcctccgc ttcgtcgggg 1020
accgcggtga gcgcttccag gcggaccgcc tcggccgggt ccagacagcg ttccgccagg 1080
cccatcactc cgctgaagct ccatgggtaa ctgcccgcgt cgcgcgcgat gttcagggag 1140
tccaccacgg cccggccgag agggccggcc cagggcaccg cgcagacgcc gagcagttgg 1200
aacgcctccg acaggccgtg tgccgctatg aaccccgcca cccagtccgc gcgctcggcg 1260
gcaggcatgg aggcgagcag tttggcccgc tcggcgaggg acacggcgcc agggccccgc 1320
gcgtcgggtg aggcgggggc gccgagcagc gctctggacc aggcgacgtc acgctggcgt 1380
acggccgcgc ggcacatgc ggcgtgcagt tcgccccgcc agtcgtcggc caccgggagc 1440
gccacgatct ccgccggggt gcggttgccg agccggggcg gccagggtggc gagcggggcc 1500
gattccacga gctggccgag ccaccaggag cgctcgcccc ggccggtggg gggcttcggg 1560
acgacgccgt cccgctccat gcccgcgtcg cactcgtgcg gcgcctcgac ggtgagggtc 1620
ggcgtgctcg atgtgtggtc gac 1643

<210> 99
 <211> 566
 <212> PRT
 <213> Streptomyces coelicolor

<400> 99

Met Asp Ala His Ala Ile Glu Glu Gly Arg Leu Arg Trp Gln Ala Arg
 1 5 10 15

Tyr Asp Ala Ala Arg Lys Arg Asp Ala Asp Phe Thr Thr Leu Ser Gly
 20 25 30

Asp Pro Val Glu Pro Val Tyr Gly Pro Arg Pro Gly Asp Glu Tyr Glu
 35 40 45

Gly Phe Glu Arg Ile Gly Trp Pro Gly Glu Tyr Pro Phe Thr Arg Gly
 50 55 60

Leu Tyr Pro Thr Gly Tyr Arg Gly Arg Thr Trp Thr Ile Arg Gln Phe
 65 70 75 80

Ala Gly Phe Gly Asn Ala Glu Gln Thr Asn Glu Arg Tyr Lys Met Ile
 85 90 95

Leu Arg Asn Gly Gly Gly Gly Leu Ser Val Ala Phe Asp Met Pro Thr
 100 105 110

Leu Met Gly Arg Asp Ser Asp Asp Pro Arg Ser Leu Gly Glu Val Gly
 115 120 125

His Cys Gly Val Ala Ile Asp Ser Ala Ala Asp Met Glu Val Leu Phe
 130 135 140

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

Pro Ala Val Pro Val Phe Cys Met Tyr Leu Val Ala Ala Glu Arg Gln
 165 170 175

Gly Val Asp Ala Ser Val Leu Asn Gly Thr Leu Gln Thr Asp Ile Phe
 180 185 190

Lys Glu Tyr Ile Ala Gln Lys Glu Trp Leu Phe Gln Pro Glu Pro His
 195 200 205

Leu Arg Leu Ile Gly Asp Leu Met Glu Tyr Cys Ala Ala Gly Ile Pro
 210 215 220

Ala Tyr Lys Pro Leu Ser Val Ser Gly Tyr His Ile Arg Glu Ala Gly
 225 230 235 240

Ala Thr Ala Ala Gln Glu Leu Ala Tyr Thr Leu Ala Asp Gly Phe Gly
 245 250 255
 Tyr Val Glu Leu Gly Leu Ser Arg Gly Leu Asp Val Asp Val Phe Ala
 260 265 270
 Pro Gly Leu Ser Phe Phe Phe Asp Ala His Leu Asp Phe Phe Glu Glu
 275 280 285
 Ile Ala Lys Phe Arg Ala Ala Arg Arg Ile Trp Ala Arg Trp Met Arg
 290 295 300
 Asp Val Tyr Gly Ala Arg Thr Asp Lys Ala Gln Trp Leu Arg Phe His
 305 310 315 320
 Thr Gln Thr Ala Gly Val Ser Leu Thr Ala Gln Gln Pro Tyr Asn Asn
 325 330 335
 Val Val Arg Thr Ala Val Glu Ala Leu Ala Ala Val Leu Gly Gly Thr
 340 345 350
 Asn Ser Leu His Thr Asn Ala Leu Asp Glu Thr Leu Ala Leu Pro Ser
 355 360 365
 Glu Gln Ala Ala Glu Ile Ala Leu Arg Thr Gln Gln Val Leu Met Glu
 370 375 380
 Glu Thr Gly Val Ala Asn Val Ala Asp Pro Leu Gly Gly Ser Trp Phe
 385 390 395 400
 Ile Glu Gln Leu Thr Asp Arg Ile Glu Ala Asp Ala Glu Lys Ile Phe
 405 410 415
 Glu Gln Ile Lys Glu Arg Gly Leu Arg Ala His Pro Asp Gly Gln His
 420 425 430
 Pro Val Gly Pro Ile Thr Ser Gly Leu Leu Arg Gly Ile Glu Asp Gly
 435 440 445
 Trp Phe Thr Gly Glu Ile Ala Glu Ser Ala Phe Arg Tyr Gln Gln Ser
 450 455 460
 Leu Glu Lys Asp Asp Lys Lys Val Val Gly Val Asn Val His Thr Gly
 465 470 475 480
 Ser Val Thr Gly Asp Leu Glu Ile Leu Arg Val Ser His Glu Val Glu
 485 490 495
 Arg Glu Gln Val Arg Val Leu Gly Glu Arg Lys Asp Ala Arg Asp Asp
 500 505 510

Ala Ala Val Arg Gly Ala Leu Asp Ala Met Leu Ala Ala Ala Arg Ser
 515 520 525

Gly Gly Asn Met Ile Gly Pro Met Leu Asp Ala Val Arg Ala Glu Ala
 530 535 540

Thr Leu Gly Glu Ile Cys Gly Val Leu Arg Asp Glu Trp Gly Val Tyr
 545 550 555 560

Thr Glu Pro Ala Gly Phe
 565

<210> 100
 <211> 1701
 <212> DNA
 <213> Streptomyces coelicolor

<400> 100
 atggacgctc atgcataga ggagggccgc cttcgctggc agggccggta cgacgcggcg 60
 cgcaagcgcg acgcggactt caccacgctc tccggagacc ccgtggagcc ggtgtacggg 120
 ccccgccccg gggacgagta cgagggcttc gagcggatcg gctggccggg cgagtacccc 180
 ttaccccgcg gcctgtatcc gaccgggtac cgggggcgta cgtggaccat ccggcagttc 240
 gccgggttcg gcaacgccga gcagaccaac gagcgctaca agatgacatc cgcacacggc 300
 ggcgggcgggc tctcggtcgc cttcgacatg ccgaccctga tgggcccgga ctccgacgac 360
 ccgcgctcgc tgggcgaggt cgggcactgc ggggtggcca tcgactcggc cgccgacatg 420
 gaagtgtgt tcaaggacat cccgctcggg gacgtgacga cctccatgac gatcagcggg 480
 cccgccgtgc ccgtgttctg catgtacctc gtcgccgccg agcgccaggg cgtcgacgca 540
 tccgtgtcga acggcacgct gcagaccgac atcttcaagg agtacatcgc ccagaaggag 600
 tggctcttcc agcccagacc ccacctccgg ctcatcgggc acctcatgga gtactgcgcg 660
 gccggcatcc ccgcctacaa gccgctctcc gtctccggct accacatccg cgaggcgggc 720
 gcgacggccg cgcaggagct ggcgtacacg ctgcgccgac gcttcggata cgtggagctg 780
 ggccctcagcc gcgggctcga cgtggacgtc ttcgcgcccg gcctctcctt cttcttcgac 840
 gcgcacctcg acttcttcga ggagatcgcc aagttccgcg cgggccgcag gatctgggcc 900
 cgctggatgc gcgacgtgta cggcgcgcg accgacaagg ccagtggtg gcggttcac 960
 acccagaccg ccggagtctc gtcaccgcg cagcagccgt acaacaacgt cgtacgcacc 1020
 gcggtggagg cgctggcggc cgtgctcggc ggcaccaact ccctgcacac caacgcgctc 1080
 gacgagaccc tcgccctgcc cagcgagcag gccgccgaga tcgccctgcg caccagcag 1140
 gtgctgatgg aggagaccgg cgtcgccaac gtcgccgacc cgctgggcgg ttcttggttc 1200
 atcgagcagc tgaccgaccg catcgaggcc gacgccgaga agatcttcga gcagatcaag 1260
 gagcgggggc tgcgcgcca ccccgacggg cagcaccgag tcggaccgat cacctccggc 1320
 ctgctgcgcg gcatcgagga cggctgggtc accggcgaga tcgccgagtc cgccttcgcg 1380

taccagcagt ccttggagaa ggacgacaag aaggtggtcg gcgtaacgt ccacaccggc 1440
 tccgtcaccg gcgacctgga gatcctgcgg gtcagccacg aggtcgagcg cgagcaggtg 1500
 cgggtcctgg gcgagcgcaa ggacgcccgg gacgacgccg ccgtgcgcgg cgccctggac 1560
 gccatgctgg ccgcggcccg ctccggcggc aacatgatcg ggccgatgct ggacgcggtg 1620
 cgcgcgaggg cgacgctggg cgagatctgc ggtgtgctgc gcgacgagtg gggggtgtac 1680
 acggaaccgg cggggttctg a 1701

<210> 101
 <211> 138
 <212> PRT
 <213> Streptomyces coelicolor

<400> 101

Met Gly Val Ala Ala Gly Pro Ile Arg Val Val Val Ala Lys Pro Gly
 1 5 10 15

Leu Asp Gly His Asp Arg Gly Ala Lys Val Ile Ala Arg Ala Leu Arg
 20 25 30

Asp Ala Gly Met Glu Val Ile Tyr Thr Gly Leu His Gln Thr Pro Glu
 35 40 45

Gln Ile Val Asp Thr Ala Ile Gln Glu Asp Ala Asp Ala Ile Gly Leu
 50 55 60

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

Leu Leu Arg Glu Arg Asp Ala Ala Asp Ile Leu Val Phe Gly Gly Gly
 85 90 95

Ile Ile Pro Glu Ala Asp Ile Ala Pro Leu Lys Glu Lys Gly Val Ala
 100 105 110

Glu Ile Phe Thr Pro Gly Ala Thr Thr Ala Ser Ile Val Asp Trp Val
 115 120 125

Arg Ala Asn Val Arg Glu Pro Ala Gly Ala
 130 135

<210> 102
 <211> 417
 <212> DNA
 <213> Streptomyces coelicolor

<400> 102

atgggtgtgg cagccggtcc gatccgcgtg gtggtggcca agccggggct cgacggccac 60

gatcgcgggg ccaaggtgat cgcgagggcc ctgcgtgacg ccggtatgga ggtgatctac 120

accgggctcc accagacgcc cgagcagatc gtcgacaccg cgatccagga ggacgccgac 180

gcgatcgggc tgtccatcct ctccggtgcg cacaacacgc tcttcgccgc cgtgatcgag 240
 ctgctccggg agcgggacgc cgcggacatc ctggtcttcg gcggcgggat catccccgag 300
 gcggacatcg ccccgctgaa ggagaagggc gtcgcggaga tcttcacgcc cggcgccacc 360
 acggcgtcca tcgtggactg ggtccgggcg aacgtgcggg agcccgcggg agcatag 417

<210> 103
 <211> 566
 <212> PRT
 <213> Streptomyces avermitilis

<400> 103

Met Asp Ala Asp Ala Ile Glu Glu Gly Arg Arg Arg Trp Gl n Ala Arg
 1 5 10 15

Tyr Asp Ala Ser Arg Lys Arg Glu Ala Asp Phe Thr Thr Leu Ser Gly
 20 25 30

Asp Pro Val Glu Pro Ala Tyr Gly Pro Arg Pro Gly Asp Ala Tyr Glu
 35 40 45

Gly Phe Glu Arg Ile Gly Trp Pro Gly Glu Tyr Pro Phe Thr Arg Gly
 50 55 60

Leu Tyr Pro Thr Gly Tyr Arg Gly Arg Thr Trp Thr Ile Arg Gl n Phe
 65 70 75 80

Ala Gly Phe Gly Asn Ala Glu Gl n Thr Asn Glu Arg Tyr Lys Lys Ile
 85 90 95

Leu Ala Asn Gly Gly Gly Gly Leu Ser Val Ala Phe Asp Met Pro Thr
 100 105 110

Leu Met Gly Arg Asp Ser Asp Asp Arg Arg Ala Leu Gly Glu Val Gly
 115 120 125

His Cys Gly Val Ala Ile Asp Ser Ala Ala Asp Met Glu Val Leu Phe
 130 135 140

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

Pro Ala Val Pro Val Phe Cys Met Tyr Leu Val Ala Ala Glu Arg Gl n
 165 170 175

Gly Val Asp Pro Ser Val Leu Asn Gly Thr Leu Gl n Thr Asp Ile Phe
 180 185 190

Lys Glu Tyr Ile Ala Gl n Lys Glu Trp Leu Phe Gl n Pro Glu Pro His
 195 200 205

Leu Arg Leu Ile Gly Asp Leu Met Glu His Cys Ala Ser Lys Ile Pro
 210 215 220
 Ala Tyr Lys Pro Leu Ser Val Ser Gly Tyr His Ile Arg Glu Ala Gly
 225 230 235 240
 Ala Thr Ala Ala Gln Glu Leu Ala Tyr Thr Leu Ala Asp Gly Phe Gly
 245 250 255
 Tyr Val Glu Leu Gly Leu Ser Arg Gly Leu Asp Val Asp Val Phe Ala
 260 265 270
 Pro Gly Leu Ser Phe Phe Phe Asp Ala His Val Asp Phe Phe Glu Glu
 275 280 285
 Ile Ala Lys Phe Arg Ala Ala Arg Arg Ile Trp Ala Arg Trp Leu Arg
 290 295 300
 Asp Val Tyr Gly Ala Lys Ser Glu Lys Ala Gln Trp Leu Arg Phe His
 305 310 315 320
 Thr Gln Thr Ala Gly Val Ser Leu Thr Ala Gln Gln Pro Tyr Asn Asn
 325 330 335
 Val Val Arg Thr Ala Val Glu Ala Leu Ala Ala Val Leu Gly Gly Thr
 340 345 350
 Asn Ser Leu His Thr Asn Ala Leu Asp Glu Thr Leu Ala Leu Pro Ser
 355 360 365
 Glu Gln Ala Ala Glu Ile Ala Leu Arg Thr Gln Gln Val Leu Met Glu
 370 375 380
 Glu Thr Gly Val Ala Asn Val Ala Asp Pro Leu Gly Gly Ser Trp Tyr
 385 390 395 400
 Val Glu Gln Leu Thr Asp Arg Ile Glu Ala Asp Ala Glu Lys Ile Phe
 405 410 415
 Glu Gln Ile Arg Glu Arg Gly Leu Arg Ala His Pro Asp Gly Arg His
 420 425 430
 Pro Ile Gly Pro Ile Thr Ser Gly Ile Leu Arg Gly Ile Glu Asp Gly
 435 440 445
 Trp Phe Thr Gly Glu Ile Ala Glu Ser Ala Phe Gln Tyr Gln Gln Ala
 450 455 460
 Leu Glu Lys Gly Asp Lys Arg Val Val Gly Val Asn Val His His Gly
 465 470 475 480

20130927_CL5646WOPCT_ST25. txt

Ser Val Thr Gly Asp Leu Glu Ile Leu Arg Val Ser His Glu Val Glu
485 490 495

Arg Glu Gln Val Arg Val Leu Gly Glu Arg Lys Ser Gly Arg Asp Asp
500 505 510

Thr Ala Val Thr Ala Ala Leu Asp Ala Met Leu Ala Ala Ala Arg Asp
515 520 525

Gly Ser Asn Met Ile Ala Pro Met Leu Asp Ala Val Arg Ala Glu Ala
530 535 540

Thr Leu Gly Glu Ile Cys Asp Val Leu Arg Glu Glu Trp Gly Val Tyr
545 550 555 560

Thr Glu Pro Ala Gly Phe
565

<210> 104
<211> 1701
<212> DNA
<213> Streptomyces avermitilis

<400> 104
tcagaaaccg gcgggctccg tgtagacccc ccactcctcc cggaggacat cgcagatctc 60
gccagcgtg gcctccgcgc ggaccgcgtc cagcatcggg gcgatcatgt tcgaccgcgc 120
gcgcgcggcg gcgagcatcg cgtccagggc cgcggttacg gccgtgtcgt cgcgccccga 180
cttccgctcg cccagcaccg gcacctgctc gcgctccacc tcgtggctga cgcgcaggat 240
ctccaggtcg cccgtcacgg acccgtggtg gacgttgacg ccgacgaccc gcttgctgcc 300
cttctccagc gcctgctggt actggaaggc cgactcggcg atctccccgg tgaaccagcc 360
gtcctcgaatg ccgcgcagga tgccggaggt gatgggcccg atcgggtgcc gcccgctccg 420
gtggggcccgc agcccgcgt ccctgatctg ttccaagatc ttctcggcgt cggcctcgaat 480
ccggtcggtc agctgtcca cgtaccagga accgcccagc ggatcggcca cgttggcgac 540
gcccgctctc tccatcagca cctgctgggt gcgcagggcg atctcggccg cctgctcgga 600
cggcagggcg agggctctct cgagggcggt ggtgtgcagc gagttcgtcc cgcagagcac 660
cgcggcgagg gcctccacgg ccgtccgtac gacgttggtg tacggctgct gcgcggtgag 720
cgagacgccc gcggtctggg tgtggaagcg cagccactgc gccttctccg acttcgcccc 780
gtacacgtcc cgcagccagc gcgcccagat gcgccgcgcc gcacggaact tggcgatctc 840
ctcgaagaag tcgacgtgcg cgtcgaagaa gaaggagagc ccgggcgcga acacgtccac 900
gtccaggccg cggctcagcc ccagctccac gtatccgaaa ccgtcggcga ggggtgtacgc 960
cagctcctgg gcggccgtgg caccggcctc ccgatgtgg taccggaga cggacagcgg 1020
cttgtagcgc gggatcttcg aggcgcagtg ctccatcagg tcgccgatga gccgcagatg 1080
gggctcgggc tggaagagcc actccttctg cgcgatgtac tccttgaaga tgtcgggtctg 1140

gaggggtgccg ttgaggacgg aggggtcgac gccctgccgc tcggccgcga ccaggtacat 1200
 gcagaagacg ggcacggcgg gcccgctgat cgtcatcgac gtcgtcacgt caccagcgg 1260
 gatgtccttg aacaggacct ccatgtcggc cgccgagtcg atcgcgaccc cgcagtgcc 1320
 gacctgccg agcgcgcggc ggtcgtcgga gtcgcgcccc atgagcgtcg gcatgtcgaa 1380
 ggccacggac agcccaccgc cgccgttggc gaggatcttc ttgtagcgct cgttggctctg 1440
 ctcggcgctt cgaacccgg cgaactgccg gatggtccag gtccggcccc ggtagccggt 1500
 cggatacaga ccgcgcgtga aggggtactc acccggccag ccgatccgct cgaaaccctc 1560
 gtacgcgtcc ccggggccgg gcccgtagc cggtccacg ggatcgccgg agagcggtgt 1620
 gaaatcggcc tcgcgcttgc gtgaggcgtc gtagcgggcc tgccagcgtc ggcggccttc 1680
 ctcgatggcg tcagcgtcca t 1701

<210> 105
 <211> 138
 <212> PRT
 <213> *Streptomyces avermitilis*

<400> 105

Met Gly Val Ala Ala Gly Pro Ile Arg Val Val Val Ala Lys Pro Gly
1 5 10 15

Leu Asp Gly His Asp Arg Gly Ala Lys Val Ile Ala Arg Ala Leu Arg
20 25 30

Asp Ala Gly Met Glu Val Ile Tyr Thr Gly Leu His Gln Thr Pro Glu
35 40 45

Gln Ile Val Gly Thr Ala Ile Gln Glu Asp Ala Asp Ala Ile Gly Leu
50 55 60

Ser Ile Leu Ser Gly Ala His Asn Thr Leu Phe Ala Ala Val Ile Asp
65 70 75 80

Leu Leu Lys Glu Arg Asp Ala Glu Asp Ile Lys Val Phe Gly Gly Gly
85 90 95

Ile Ile Pro Glu Ala Asp Ile Ala Pro Leu Lys Glu Lys Gly Val Ala
100 105 110

Glu Ile Phe Thr Pro Gly Ala Thr Thr Ala Ser Ile Val Glu Trp Val
115 120 125

Arg Ala Asn Val Arg Gln Pro Ala Gly Ala
130 135

<210> 106
 <211> 1701
 <212> DNA

<213> Streptomyces avermitilis

<400> 106

tcagaaaccg gcgggctccg tgtagacccc ccactcctcc cggaggacat cgcagatctc	60
gcccagcgtg gcctccgcgc ggaccgcgtc cagcatcggg gcgatcatgt tcgacccgtc	120
gcgcgcggcg gcgagcatcg cgtccagggc cgcggttacg gccgtgtcgt cgcgccccga	180
cttccgctcg cccagcaccg gcacctgctc gcgctccacc tcgtggctga cgcgcaggat	240
ctccaggctg cccgtcacgg acccgtgggt gacgttgacg ccgacgaccc gcttgctgcc	300
cttctccagc gcctgctggt actggaaggc cgactcggcg atctccccgg tgaaccagcc	360
gtcctcgatg ccgcgcagga tgccggaggt gatgggcccg atcgggtgcc gcccgctccg	420
gtgggcccgc agcccgcgct ccctgatctg ttcgaagatc ttctcggcgt cggcctcgat	480
ccggtcggtc agctgctcca cgtaccagga accgcccagc ggatcggcca cgttggcgac	540
gcccgctctc tccatcagca cctgctgggt gcgcagggcg atctcggccg cctgctcgga	600
cggcagggcg agggctctgt cgagggcggt ggtgtgcagc gagttcgtcc cgccgagcac	660
cgcggcgagg gcctccacgg ccgtccgtac gacgttggtg tacggctgct gcgcggtgag	720
cgagacgccc gcggtctggg tgtggaagcg cagccactgc gccttctccg acttcgcccc	780
gtacacgtcc cgcagccagc gcgcccagat gcgccgcgcc gcacggaact tggcgatctc	840
ctcgaagaag tcgacgtgcg cgtcgaagaa gaaggagagc ccgggcgcgga acacgtccac	900
gtccaggccg cggctcagcc ccagctccac gtatccgaaa ccgtcggcga ggggtgtacgc	960
cagctcctgg gcggccgtgg caccggcctc ccggatgtgg taccggaga cggacagcgg	1020
cttgtagcgc gggatcttcg aggcgcagtg ctccatcagg tcgccgatga gccgcagatg	1080
gggctcgggc tggaagagcc actccttctg cgcgatgtac tccttgaaga tgtcggctctg	1140
gagggtgccg ttgaggacgg aggggtcgac gccctgccgc tcggccgcga ccaggtagat	1200
gcagaagacg ggcacggcgg gcccgctgat cgtcatcgac gtcgtcacgt caccagcgg	1260
gatgtccttg aacaggacct ccatgtcggc cgccgagtcg atcgcgaccc cgcagtgcgc	1320
gacctcgccg agcgcgcggc ggtcgtcgga gtcgcgcccc atgagcgtcg gcatgtcgaa	1380
ggccacggac agcccaccgc cgccgttggc gaggatcttc ttgtagcgct cgttggtctg	1440
ctcggcggtg ccgaaccggc cgaactgccg gatggtccag gtccggcccc ggtagccggt	1500
cggatacaga ccgcgcgtga aggggtactc acccgccag ccgatccgct cgaaaccctc	1560
gtacgcgtcc ccgggcccgg gcccgtagc cggctccacg ggatcgccgg agagcgtggt	1620
gaaatcggcc tcgcgcttgc gtgaggcgtc gtagcgggcc tgccagcgtc ggcggccttc	1680
ctcgatggcg tcagcgtcca t	1701

<210> 107

<211> 267

<212> PRT

<213> Saccharomyces cerevisiae

<400> 107

20130927_CL5646WOPCT_ST25. txt

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

Leu Ile Thr Gly Ala Ser Ala Gly Ile Gly Lys Ala Thr Ala Leu Glu
20 25 30

Tyr Leu Glu Ala Ser Asn Gly Asp Met Lys Leu Ile Leu Ala Ala Arg
35 40 45

Arg Leu Glu Lys Leu Glu Glu Leu Lys Lys Thr Ile Asp Gln Glu Phe
50 55 60

Pro Asn Ala Lys Val His Val Ala Gln Leu Asp Ile Thr Gln Ala Glu
65 70 75 80

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

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

Gly Gln Ile Ala Thr Glu Asp Ile Gln Asp Val Phe Asp Thr Asn Val
115 120 125

Thr Ala Leu Ile Asn Ile Thr Gln Ala Val Leu Pro Ile Phe Gln Ala
130 135 140

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

Ala Tyr Pro Thr Gly Ser Ile Tyr Cys Ala Ser Lys Phe Ala Val Gly
165 170 175

Ala Phe Thr Asp Ser Leu Arg Lys Glu Leu Ile Asn Thr Lys Ile Arg
180 185 190

Val Ile Leu Ile Ala Pro Gly Leu Val Glu Thr Glu Phe Ser Leu Val
195 200 205

Arg Tyr Arg Gly Asn Glu Glu Gln Ala Lys Asn Val Tyr Lys Asp Thr
210 215 220

Thr Pro Leu Met Ala Asp Asp Val Ala Asp Leu Ile Val Tyr Ala Thr
225 230 235 240

Ser Arg Lys Gln Asn Thr Val Ile Ala Asp Thr Leu Ile Phe Pro Thr
245 250 255

Asn Gln Ala Ser Pro His His Ile Phe Arg Gly
260 265

<210> 108
 <211> 804
 <212> DNA
 <213> Saccharomyces cerevisiae

<400> 108
 atgtccaag gtagaaaagc tgcagaaaga ttggctaaga agactgtcct cattacaggt 60
 gcatctgctg gtattggtaa ggcgaccgca ttagagtact tggaggcatc caatggtgat 120
 atgaaactga tcttggctgc tagaagatta gaaaagctcg aggaattgaa gaagaccatt 180
 gatcaagagt ttccaaacgc aaaagttcat gtggcccagc tggatatcac tcaagcagaa 240
 aaaatcaagc ccttcattga aaacttgcca caagagttca aggatattga cattctggtg 300
 aacaatgccg gaaaggctct tggcagtgac cgtgtggggc agatcgcaac ggaggatata 360
 caggacgtgt ttgacaccaa cgtcacggct ttaatcaata tcacacaagc tgtactgccc 420
 atattccaag ccaagaattc aggagatatt gtaaatttgg gttcaatcgc tggcagagac 480
 gcatacccaa caggttctat ctattgtgcc tctaagtttg ccgtgggggc gttcactgat 540
 agtttgagaa aggagctcat caacactaaa attagagtca ttctaattgc accagggcta 600
 gtcgagactg aattttcact agttagatac agaggtaacg aggaacaagc caagaatggt 660
 tacaaggata ctacccatt gatggctgat gacgtggctg atctgacgt ctatgcaact 720
 tccagaaaac aaaatactgt aattgcagac actttaatct ttccaacaaa ccaagcgtca 780
 cctcatcata tcttccgtgg ataa 804

<210> 109
 <211> 139
 <212> PRT
 <213> Saccharomyces cerevisiae

<400> 109
 Met Ser Glu Ile Thr Leu Gly Lys Tyr Leu Phe Glu Arg Leu Lys Gln
 1 5 10 15
 Val Asn Val Asn Thr Val Phe Gly Leu Pro Gly Asp Phe Asn Leu Ser
 20 25 30
 Leu Leu Asp Lys Ile Tyr Glu Val Glu Gly Met Arg Trp Ala Gly Asn
 35 40 45
 Ala Asn Glu Leu Asn Ala Ala Tyr Ala Ala Asp Gly Tyr Ala Arg Ile
 50 55 60
 Lys Gly Met Ser Cys Ile Ile Thr Thr Phe Gly Val Gly Glu Leu Ser
 65 70 75 80
 Ala Leu Asn Gly Ile Ala Gly Ser Tyr Ala Glu His Val Gly Val Leu
 85 90 95
 His Val Val Gly Val Pro Ser Ile Ser Ala Gln Ala Lys Gln Leu Leu

100

105

110

Leu His His Thr Leu Gly Asn Gly Asp Phe Thr Val Phe His Arg Met
 115 120 125

Ser Ala Asn Ile Ser Glu Thr Thr Ala Met Ile
 130 135

<210> 110
 <211> 1689
 <212> DNA
 <213> Saccharomyces cerevisiae

<400> 110
 atgtctgaaa ttactttggg taaatatattg ttcgaaagat taaagcaagt caacgttaac 60
 accgttttcg gtttgccagg tgacttcaac ttgtccttgt tggacaagat ctacgaagtt 120
 gaaggatga gatgggctgg taacgccaac gaattgaacg ctgcttacgc cgctgatggt 180
 tacgctcgta tcaagggtat gtcttgatc atcaccacct tcggtgtcgg tgaattgtct 240
 gctttgaacg gtattgccgg ttcttacgct gaacacgtcg gtgttttgca cgttgttggt 300
 gtcccatcca tctctgctca agctaagcaa ttgtttgttg accacacctt gggtaacggt 360
 gacttcactg tttccacag aatgtctgcc aacatttctg aaaccactgc tatgatcact 420
 gacattgcta ccgccccagc tgaaattgac agatgtatca gaaccactta cgtcacccaa 480
 agaccagtct acttaggttt gccagctaac ttggctgact tgaacgtccc agctaagttg 540
 ttgcaaactc caattgacat gtctttgaag ccaaacgatg ctgaatccga aaaggaagtc 600
 attgacacca tcttggtttt ggtcaaggat gctaagaacc cagttatctt ggctgatgct 660
 tgttgttcca gacacgacgt caaggctgaa actaagaagt tgattgactt gactcaattc 720
 ccagctttcg tcacccaat gggtaagggt tccattgacg aacaacaccc aagatacggg 780
 ggtgtttacg tcggtacctt gtccaagcca gaagttaagg aagccgttga atctgctgac 840
 ttgattttgt ctgtcggcgtc ttgtttgtct gatttcaaca ccggttcttt ctcttactct 900
 tacaagacca agaacattgt cgaattccac tccgaccaca tgaagatcag aaacgccact 960
 ttcccagggtg tccaaatgaa attcgttttg caaaagttgt tgaccactat tgctgacgcc 1020
 gctaagggtt acaagccagt tgctgtccca gctagaactc cagctaacgc tgctgtccca 1080
 gcttctaccc cattgaagca agaattgatg tggaaccaat tgggtaactt cttgcaagaa 1140
 ggtgatgttg tcattgctga aaccggtacc tccgctttcg gtatcaacca aaccactttc 1200
 ccaaacaaca cctacggtat ctctcaagtc ttatgggggtt ccattgggtt caccactggt 1260
 gctaccttgg gtgctgcttt cgctgctgaa gaaattgatc caaagaagag agttatctta 1320
 ttcatgggtg acggttcttt gcaattgact gttcaagaaa tctccaccat gatcagatgg 1380
 ggcttgaagc catacttggt cgtcttgaac aacgatgggt acaccattga aaagttgatt 1440
 cacggtccaa aggctcaata caacgaaatt caaggttggg accacctatc cttgttgcca 1500
 actttcgggtg ctaaggacta tgaaaccac agagtcgcta ccaccggtga atgggacaag 1560

ttgacccaag acaagtcttt caacgacaac tctaagatca gaatgattga aatcatgttg 1620
ccagtcttcg atgtccaca aaacttgggt gaacaagcta agttgactgc tgctaccaac 1680
gctaagcaa 1689

<210> 111
<211> 563
<212> PRT
<213> Saccharomyces cerevisiae

<400> 111

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

Val Asn Cys Asn Thr Val Phe Gly Leu Pro Gly Asp Phe Asn Leu Ser
20 25 30

Leu Leu Asp Lys Leu Tyr Glu Val Lys Gly Met Arg Trp Ala Gly Asn
35 40 45

Ala Asn Glu Leu Asn Ala Ala Tyr Ala Ala Asp Gly Tyr Ala Arg Ile
50 55 60

Lys Gly Met Ser Cys Ile Ile Thr Thr Phe Gly Val Gly Glu Leu Ser
65 70 75 80

Ala Leu Asn Gly Ile Ala Gly Ser Tyr Ala Glu His Val Gly Val Leu
85 90 95

His Val Val Gly Val Pro Ser Ile Ser Ser Gln Ala Lys Gln Leu Leu
100 105 110

Leu His His Thr Leu Gly Asn Gly Asp Phe Thr Val Phe His Arg Met
115 120 125

Ser Ala Asn Ile Ser Glu Thr Thr Ala Met Ile Thr Asp Ile Ala Asn
130 135 140

Ala Pro Ala Glu Ile Asp Arg Cys Ile Arg Thr Thr Tyr Thr Thr Gln
145 150 155 160

Arg Pro Val Tyr Leu Gly Leu Pro Ala Asn Leu Val Asp Leu Asn Val
165 170 175

Pro Ala Lys Leu Leu Glu Thr Pro Ile Asp Leu Ser Leu Lys Pro Asn
180 185 190

Asp Ala Glu Ala Glu Ala Glu Val Val Arg Thr Val Val Glu Leu Ile
195 200 205

Lys Asp Ala Lys Asn Pro Val Ile Leu Ala Asp Ala Cys Ala Ser Arg
Page 159

210

215

220

His Asp Val Lys Ala Glu Thr Lys Lys Leu Met Asp Leu Thr Gl n Phe
 225 230 235 240

Pro Val Tyr Val Thr Pro Met Gly Lys Gly Ala Ile Asp Gl u Gl n His
 245 250 255

Pro Arg Tyr Gly Gly Val Tyr Val Gly Thr Leu Ser Arg Pro Gl u Val
 260 265 270

Lys Lys Ala Val Gl u Ser Ala Asp Leu Ile Leu Ser Ile Gly Ala Leu
 275 280 285

Leu Ser Asp Phe Asn Thr Gly Ser Phe Ser Tyr Ser Tyr Lys Thr Lys
 290 295 300

Asn Ile Val Gl u Phe His Ser Asp His Ile Lys Ile Arg Asn Ala Thr
 305 310 315 320

Phe Pro Gly Val Gl n Met Lys Phe Ala Leu Gl n Lys Leu Leu Asp Ala
 325 330 335

Ile Pro Gl u Val Val Lys Asp Tyr Lys Pro Val Ala Val Pro Ala Arg
 340 345 350

Val Pro Ile Thr Lys Ser Thr Pro Ala Asn Thr Pro Met Lys Gl n Gl u
 355 360 365

Trp Met Trp Asn His Leu Gly Asn Phe Leu Arg Gl u Gly Asp Ile Val
 370 375 380

Ile Ala Gl u Thr Gly Thr Ser Ala Phe Gly Ile Asn Gl n Thr Thr Phe
 385 390 395 400

Pro Thr Asp Val Tyr Ala Ile Val Gl n Val Leu Trp Gly Ser Ile Gly
 405 410 415

Phe Thr Val Gly Ala Leu Leu Gly Ala Thr Met Ala Ala Gl u Gl u Leu
 420 425 430

Asp Pro Lys Lys Arg Val Ile Leu Phe Ile Gly Asp Gly Ser Leu Gl n
 435 440 445

Leu Thr Val Gl n Gl u Ile Ser Thr Met Ile Arg Trp Gly Leu Lys Pro
 450 455 460

Tyr Ile Phe Val Leu Asn Asn Asn Gly Tyr Thr Ile Gl u Lys Leu Ile
 465 470 475 480

His Gly Pro His Ala Gl u Tyr Asn Gl u Ile Gl n Gly Trp Asp His Leu
 Page 160

485

490

495

Ala Leu Leu Pro Thr Phe Gly Ala Arg Asn Tyr Glu Thr His Arg Val
 500 505 510

Ala Thr Thr Gly Glu Trp Glu Lys Leu Thr Gln Asp Lys Asp Phe Gln
 515 520 525

Asp Asn Ser Lys Ile Arg Met Ile Glu Val Met Leu Pro Val Phe Asp
 530 535 540

Ala Pro Gln Asn Leu Val Lys Gln Ala Gln Leu Thr Ala Ala Thr Asn
 545 550 555 560

Ala Lys Gln

<210> 112
 <211> 1689
 <212> DNA
 <213> Saccharomyces cerevisiae

<400> 112
 atgtctgaaa taaccttagg taaatattta tttgaaagat tgagccaagt caactgtaac 60
 accgtcttcg gtttgccagg tgactttaac ttgtctcttt tggataagct ttatgaagtc 120
 aaaggatga gatgggctgg taacgctaac gaattgaacg ctgcctatgc tgctgatggt 180
 tacgctcgta tcaagggat gtcctgtatt attaccacct tcggtgttgg tgaattgtct 240
 gctttgaatg gtattgccgg ttcttacgct gaacatgtcg gtgttttgca cgttgttggt 300
 gttccatcca tctcttctca agctaagcaa ttgttgttgc atcatacctt gggtaacggt 360
 gacttcactg tttccacag aatgtctgcc aacatttctg aaaccactgc catgatcact 420
 gatattgcta acgctccagc tgaaattgac agatgtatca gaaccaccta cactacccaa 480
 agaccagtct acttgggttt gccagctaac ttggttgact tgaacgtccc agccaagtta 540
 ttggaaactc caattgactt gtctttgaag ccaaacgacg ctgaagctga agctgaagtt 600
 gttagaactg ttgttgaatt gatcaaggat gctaagaacc cagttatctt ggctgatgct 660
 tgtgttcta gacatgatgt caaggctgaa actaagaagt tgatggactt gactcaattc 720
 ccagtttacg tcacccaat gggttaagggt gctattgacg aacaacaccc aagatacggg 780
 ggtgtttacg ttggtacctt gtctagacca gaagttaaga aggctgtaga atctgctgat 840
 ttgatattgt ctatcggtgc tttgttgtct gatttcaata ccggttcttt ctcttactcc 900
 tacaagacca aaaatatcgt tgaattccac tctgaccaca tcaagatcag aaacgccacc 960
 ttcccagggtg ttcaaataaa atttgccttg caaaaattgt tggatgctat tccagaagtc 1020
 gtcaaggact acaaacctgt tgctgtccca gctagagttc caattaccaa gtctactcca 1080
 gctaacactc caatgaagca agaattgatg tggaaccatt tgggtaactt cttgagagaa 1140
 ggtgatattg ttattgctga aaccggtact tccgccttcg gtattaacca aactactttc 1200

```

ccaacagatg tatacgctat cgtccaagtc ttgtgggggtt ccattgggtt cacagtcggc      1260
gctctattgg gtgctactat ggccgctgaa gaacttgatc caaagaagag agttatttta      1320
ttcattgggtg acggttctct acaattgact gttcaagaaa tctctacat gattagatgg      1380
ggtttgaagc catacatttt tgtcttgaat aacaacgggtt acaccattga aaaattgatt      1440
cacggtcctc atgccgaata taatgaaatt caaggttggg accacttggc cttattgcca      1500
acttttgggtg ctagaaacta cgaaaccac agagttgcta ccactgggtga atgggaaaag      1560
ttgactcaag acaaggactt ccaagacaac tctaagatta gaatgattga agttatgttg      1620
ccagtctttg atgctccaca aaacttgggtt aaacaagctc aattgactgc cgctactaac      1680
gctaaacaa                                                                    1689

```

```

<210> 113
<211> 533
<212> PRT
<213> Saccharomyces cerevisiae

```

```
<400> 113
```

```
Met Ser Glu Ile Thr Leu Gly Lys Tyr Leu Phe Glu Arg Leu Lys Gln
1           5           10           15
```

```
Val Asn Val Asn Thr Ile Phe Gly Leu Pro Gly Asp Phe Asn Leu Ser
          20           25           30
```

```
Leu Leu Asp Lys Ile Tyr Glu Val Asp Gly Leu Arg Trp Ala Gly Asn
          35           40           45
```

```
Ala Asn Glu Leu Asn Ala Ala Tyr Ala Ala Asp Gly Tyr Ala Arg Ile
          50           55           60
```

```
Lys Gly Leu Ser Val Leu Val Thr Thr Phe Gly Val Gly Glu Leu Ser
65           70           75           80
```

```
Ala Leu Asn Gly Ile Ala Gly Ser Tyr Ala Glu His Val Gly Val Leu
          85           90           95
```

```
His Val Val Gly Val Pro Ser Ile Ser Ala Gln Ala Lys Gln Leu Leu
          100          105          110
```

```
Leu His His Thr Leu Gly Asn Gly Asp Phe Thr Val Phe His Arg Met
          115          120          125
```

```
Ser Ala Asn Ile Ser Glu Thr Thr Ser Met Ile Thr Asp Ile Ala Thr
          130          135          140
```

```
Ala Pro Ser Glu Ile Asp Arg Leu Ile Arg Thr Thr Phe Ile Thr Gln
          145          150          155          160
```

```
Arg Pro Ser Tyr Leu Gly Leu Pro Ala Asn Leu Val Asp Leu Lys Val
Page 162

```

165

170

175

Pro Gly Ser Leu₁₈₀ Leu Glu Lys Pro Ile₁₈₅ Asp Leu Ser Leu Lys₁₉₀ Pro Asn
 Asp Pro Glu₁₉₅ Ala Glu Lys Glu Val₂₀₀ Ile Asp Thr Val Leu₂₀₅ Glu Leu Ile
 Gln Asn₂₁₀ Ser Lys Asn Pro Val₂₁₅ Ile Leu Ser Asp Ala₂₂₀ Cys Ala Ser Arg
 His₂₂₅ Asn Val Lys Lys Glu₂₃₀ Thr Gln Lys Leu Ile₂₃₅ Asp Leu Thr Gln Phe₂₄₀
 Pro Ala Phe Val Thr₂₄₅ Pro Leu Gly Lys Gly₂₅₀ Ser Ile Asp Glu Gln His₂₅₅
 Pro Arg Tyr Gly₂₆₀ Gly Val Tyr Val Gly₂₆₅ Thr Leu Ser Lys Gln₂₇₀ Asp Val
 Lys Gln Ala₂₇₅ Val Glu Ser Ala Asp₂₈₀ Leu Ile Leu Ser Val₂₈₅ Gly Ala Leu
 Leu Ser₂₉₀ Asp Phe Asn Thr Gly₂₉₅ Ser Phe Ser Tyr Ser₃₀₀ Tyr Lys Thr Lys
 Asn₃₀₅ Val Val Glu Phe His₃₁₀ Ser Asp Tyr Val Lys₃₁₅ Val Lys Asn Ala Thr₃₂₀
 Phe Leu Gly Val Gln₃₂₅ Met Lys Phe Ala Leu₃₃₀ Gln Asn Leu Leu Lys₃₃₅ Val
 Ile Pro Asp Val₃₄₀ Val Lys Gly Tyr Lys₃₄₅ Ser Val Pro Val Pro₃₅₀ Thr Lys
 Thr Pro Ala₃₅₅ Asn Lys Gly Val Pro₃₆₀ Ala Ser Thr Pro Leu₃₆₅ Lys Gln Glu
 Trp Leu₃₇₀ Trp Asn Glu Leu Ser₃₇₅ Lys Phe Leu Gln Glu₃₈₀ Gly Asp Val Ile
 Ile₃₈₅ Ser Glu Thr Gly Thr₃₉₀ Ser Ala Phe Gly Ile₃₉₅ Asn Gln Thr Ile Phe₄₀₀
 Pro Lys Asp Ala Tyr₄₀₅ Gly Ile Ser Gln Val₄₁₀ Leu Trp Gly Ser Ile₄₁₅ Gly
 Phe Thr Thr Gly₄₂₀ Ala Thr Leu Gly Ala₄₂₅ Ala Phe Ala Ala Glu₄₃₀ Glu Ile
 Asp Pro Asn Lys Arg Val Ile Leu Phe Ile Gly Asp Gly Ser Leu Gln

435

440

445

Leu Thr Val Gln Glu Ile Ser Thr Met Ile Arg Trp Gly Leu Lys Pro
 450 455 460

Tyr Leu Phe Val Leu Asn Asn Asp Gly Tyr Thr Ile Glu Lys Leu Ile
 465 470 475 480

His Gly Pro His Ala Glu Tyr Asn Glu Ile Gln Thr Trp Asp His Leu
 485 490 495

Ala Leu Leu Pro Ala Phe Gly Ala Lys Lys Tyr Glu Asn His Lys Ile
 500 505 510

Ala Thr Thr Gly Glu Trp Asp Ala Leu Thr Thr Asp Ser Glu Phe Gln
 515 520 525

Lys Asn Ser Val Ile
 530

<210> 114
 <211> 1599
 <212> DNA
 <213> Saccharomyces cerevisiae

<400> 114
 atgtctgaaa ttactcttgg aaaatactta tttgaaagat tgaagcaagt taatgttaac 60
 accatTTTTg ggctaccagg cgacttcaac ttgtccctat tggacaagat ttacgaggta 120
 gatggattga gatgggctgg taatgcaaat gagctgaacg cgcctatgc cgccgatggt 180
 tacgcacgca tcaagggttt atctgtgctg gtaactactt ttggcgtagg tgaattatcc 240
 gccttgaatg gtattgcagg atcgtatgca gaacacgtcg gtgtactgca tgttgttgggt 300
 gtcccctcta tctccgctca ggctaagcaa ttgttgttgc atcatacctt gggtaacggt 360
 gattttaccg tttttcacag aatgtccgcc aatatctcag aaactacatc aatgattaca 420
 gacattgcta cagccccttc agaaatcgat aggttgatca ggacaacatt tataacacaa 480
 aggcctagct acttgggggtt gccagcgaat ttggtagatc taaaggttcc tggttctctt 540
 ttggaaaaac cgattgatct atcattaaaa cctaacgatc ccgaagctga aaaggaagtt 600
 attgataccg tactagaatt gatccagaat tcgaaaaacc ctgttatact atcggatgcc 660
 tgtgcttcta ggcacaacgt taaaaaagaa acccagaagt taattgattt gacgcaattc 720
 ccagcttttg tgacacctct aggtaaaggg tcaatagatg aacagcatcc cagatatggc 780
 ggtgtttatg tgggaacgct gtccaaacaa gacgtgaaac aggccgttga gtcggctgat 840
 ttgatccttt cggtcggtgc tttgctctct gatTTtaaca caggttcgtt ttctactcc 900
 tacaagacta aaaatgtagt ggagtttcat tccgattacg taaaggtgaa gaacgctacg 960
 ttctcgggtg tacaatgaa atttgacta caaaacttac tgaaggttat tcccgatgtt 1020
 gttaagggct acaagagcgt tcccgtagca accaaaactc ccgcaaacaa aggtgtacct 1080

```

gctagcacgc ccttgaaaca agagtgggtg tggaacgaat tgtccaaatt cttgcaagaa      1140
ggtagatgta tcatttccga gaccggcacg tctgccttcg gtatcaatca aactatcttt      1200
cctaaggacg cctacggtat ctgcgagggt ttgtgggggt ccatcggttt tacaacagga      1260
gcaactttag gtgctgcctt tgccgctgag gagattgacc ccaacaagag agtcatctta      1320
ttcatagggtg acgggtcttt gcagttaacc gtccaagaaa tctccaccat gatcagatgg      1380
gggttaaagc cgtatctttt tgtccttaac aacgacgggt acactatcga aaagctgatt      1440
catgggcctc acgcagagta caacgaaatc cagacctggg atcacctcgc cctgttgccc      1500
gcatttggtg cgaaaaagta cgaaaatcac aagatcgcca ctacgggtga gtgggatgcc      1560
ttaaccactg attcagagtt ccagaaaaac tcggtgatc                                1599

```

<210> 115
 <211> 564
 <212> PRT
 <213> Candi da glabrata

<400> 115

```

Met Ser Glu Ile Thr Leu Gly Arg Tyr Leu Phe Glu Arg Leu Asn Gln
1           5           10           15

```

```

Val Asp Val Lys Thr Ile Phe Gly Leu Pro Gly Asp Phe Asn Leu Ser
           20           25           30

```

```

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

```

```

Ala Asn Glu Leu Asn Ala Ala Tyr Ala Ala Asp Gly Tyr Ala Arg Ile
           50           55           60

```

```

Lys Gly Met Ser Cys Ile Ile Thr Thr Phe Gly Val Gly Glu Leu Ser
65           70           75           80

```

```

Ala Leu Asn Gly Ile Ala Gly Ser Tyr Ala Glu His Val Gly Val Leu
           85           90           95

```

```

His Val Val Gly Val Pro Ser Ile Ser Ser Gln Ala Lys Gln Leu Leu
           100          105          110

```

```

Leu His His Thr Leu Gly Asn Gly Asp Phe Thr Val Phe His Arg Met
           115          120          125

```

```

Ser Ala Asn Ile Ser Glu Thr Thr Ala Met Val Thr Asp Ile Ala Thr
           130          135          140

```

```

Ala Pro Ala Glu Ile Asp Arg Cys Ile Arg Thr Thr Tyr Ile Thr Gln
145          150          155          160

```

```

Arg Pro Val Tyr Leu Gly Leu Pro Ala Asn Leu Val Asp Leu Lys Val

```

165

170

175

Pro Ala Lys Leu₁₈₀ Leu Glu Thr Pro Ile₁₈₅ Asp Leu Ser Leu Lys₁₉₀ Pro Asn
 Asp Pro Glu₁₉₅ Ala Glu Thr Glu Val₂₀₀ Val Asp Thr Val Leu₂₀₅ Glu Leu Ile
 Lys Ala₂₁₀ Ala Lys Asn Pro Val₂₁₅ Ile Leu Ala Asp Ala₂₂₀ Cys Ala Ser Arg
 His₂₂₅ Asp Val Lys Ala Glu₂₃₀ Thr Lys Lys Leu Ile₂₃₅ Asp Ala Thr Gln Phe₂₄₀
 Pro Ser Phe Val Thr₂₄₅ Pro Met Gly Lys Gly₂₅₀ Ser Ile Asp Glu Gln His₂₅₅
 Pro Arg Phe Gly₂₆₀ Gly Val Tyr Val Gly₂₆₅ Thr Leu Ser Arg Pro₂₇₀ Glu Val
 Lys Glu Ala₂₇₅ Val Glu Ser Ala Asp₂₈₀ Leu Ile Leu Ser Val₂₈₅ Gly Ala Leu
 Leu Ser₂₉₀ Asp Phe Asn Thr Gly₂₉₅ Ser Phe Ser Tyr Ser₃₀₀ Tyr Lys Thr Lys
 Asn Ile₃₀₅ Val Glu Phe His₃₁₀ Ser Asp Tyr Ile Lys₃₁₅ Ile Arg Asn Ala Thr₃₂₀
 Phe Pro Gly Val Gln₃₂₅ Met Lys Phe Ala Leu₃₃₀ Gln Lys Leu Leu Asn Ala₃₃₅
 Val Pro Glu Ala₃₄₀ Ile Lys Gly Tyr Lys₃₄₅ Pro Val Pro Val Pro₃₅₀ Ala Arg
 Val Pro Glu₃₅₅ Asn Lys Ser Cys Asp₃₆₀ Pro Ala Thr Pro Leu₃₆₅ Lys Gln Glu
 Trp Met₃₇₀ Trp Asn Gln Val Ser₃₇₅ Lys Phe Leu Gln Glu₃₈₀ Gly Asp Val Val
 Ile Thr Glu Thr Gly Thr₃₉₀ Ser Ala Phe Gly Ile₃₉₅ Asn Gln Thr Pro Phe₄₀₀
 Pro Asn Asn Ala Tyr₄₀₅ Gly Ile Ser Gln Val₄₁₀ Leu Trp Gly Ser Ile₄₁₅ Gly
 Phe Thr Thr Gly₄₂₀ Ala Cys Leu Gly Ala₄₂₅ Ala Phe Ala Ala Glu Glu Ile₄₃₀
 Asp Pro Lys Lys Arg Val Ile Leu Phe Ile Gly Asp Gly Ser Leu Gln

435

440

445

Leu Thr Val Gln Glu Ile Ser Thr Met Ile Arg Trp Gly Leu Lys Pro
 450 455 460

Tyr Leu Phe Val Leu Asn Asn Asp Gly Tyr Thr Ile Glu Arg Leu Ile
 465 470 475 480

His Gly Glu Lys Ala Gly Tyr Asn Asp Ile Gln Asn Trp Asp His Leu
 485 490 495

Ala Leu Leu Pro Thr Phe Gly Ala Lys Asp Tyr Glu Asn His Arg Val
 500 505 510

Ala Thr Thr Gly Glu Trp Asp Lys Leu Thr Gln Asp Lys Glu Phe Asn
 515 520 525

Lys Asn Ser Lys Ile Arg Met Ile Glu Val Met Leu Pro Val Met Asp
 530 535 540

Ala Pro Thr Ser Leu Ile Glu Gln Ala Lys Leu Thr Ala Ser Ile Asn
 545 550 555 560

Ala Lys Gln Glu

<210> 116
 <211> 1692
 <212> DNA
 <213> Candi da gl abrata

<400> 116
 atgtctgaga ttactttggg tagatacttg ttcgagagat tgaaccaagt cgacgttaag 60
 accatcttcg gtttgccagg tgacttcaac ttgtccctat tggacaagat ctacgaagtt 120
 gaaggatga gatgggctgg taacgctaac gaattgaacg ctgcttacgc tgctgacggt 180
 tacgctagaa tcaagggat gtcctgtatc atcaccacct tcggtgtcgg tgaattgtct 240
 gccttgaacg gtattgccgg ttcttacgct gaacacgtcg gtgtcttgca cgtcgtcgggt 300
 gtcccatcca tctcctctca agctaagcaa ttgttgttgc accacacctt gggtaacggt 360
 gacttcaactg tcttccacag aatgtccgct aacatctctg agaccaccgc tatgggtcact 420
 gacatcgcta ccgctccagc tgagatcgac agatgtatca gaaccaccta catcacccaa 480
 agaccagtct acttgggtct accagctaac ttgggtcgacc taaaggtccc agccaagctt 540
 ttggaaaccc caattgactt gtccttgaag ccaaacgacc cagaagccga aactgaagtc 600
 gttgacaccg tcttgggaatt gatcaaggct gctaagaacc cagttatctt ggctgatgct 660
 tgtgcttcca gacacgacgt caaggctgaa accaagaagt tgattgacgc cactcaattc 720
 ccatccttcg ttaccccaat gggttaagggt tccatcgacg aacaacaccc aagattcgggt 780
 ggtgtctacg tcggtacctt gtccagacca gaagttaagg aagctgttga atccgctgac 840

```

ttgatcttgt ctgtcggtagc tttgttgtcc gatttcaaca ctggttcttt ctcttactct      900
tacaagacca agaacatcgt cgaattccac tctgactaca tcaagatcag aaacgctacc      960
ttcccagggtg tccaaatgaa gttcgcctttg caaaagttgt tgaacgccgt cccagaagct    1020
atcaaggggtt acaagccagt ccctgtccca gctagagtcc cagaaaacaa gtcctgtgac    1080
ccagctaccc cattgaagca agaatggatg tggaaccaag ttccaagtt cttgcaagaa    1140
ggtgatgttg ttatcactga aaccggtacc tccgcttttg gtatcaacca aaccccatc      1200
ccaaacaacg cttacggtat ctccaagtt ctatgggggtt ccatcggttt caccaccggt      1260
gcttgtttgg gtgccgcttt cgctgtgaa gaaatcgacc caaagaagag agttatcttg      1320
ttcattggtg acggttcttt gcaattgact gtccaagaaa tctccaccat gatcagatgg      1380
ggcttgaagc catacttggt cgtcttgaac aacgacggtt acaccatcga aagattgatt      1440
cacggtgaaa aggtctggtta caacgacatc caaaactggg accacttggc tctattgcc      1500
accttcggtg ctaaggacta cgaaaaccac agagtcgccca ccaccggtga atgggacaag      1560
ttgaccaag acaaggaatt caacaagaac tccaagatca gaatgatcga agttatgttg      1620
ccagttatgg acgctccaac ttccttgatt gaacaagcta agttgaccgc ttccatcaac      1680
gctaagcaag aa                                                              1692

```

```

<210> 117
<211> 596
<212> PRT
<213> Pichia stipitis

```

```
<400> 117
```

```
Met Ala Glu Val Ser Leu Gly Arg Tyr Leu Phe Glu Arg Leu Tyr Gln
1          5          10          15

```

```
Leu Gln Val Gln Thr Ile Phe Gly Val Pro Gly Asp Phe Asn Leu Ser
20          25          30

```

```
Leu Leu Asp Lys Ile Tyr Glu Val Glu Asp Ala His Gly Lys Asn Ser
35          40          45

```

```
Phe Arg Trp Ala Gly Asn Ala Asn Glu Leu Asn Ala Ser Tyr Ala Ala
50          55          60

```

```
Asp Gly Tyr Ser Arg Val Lys Arg Leu Gly Cys Leu Val Thr Thr Phe
65          70          75          80

```

```
Gly Val Gly Glu Leu Ser Ala Leu Asn Gly Ile Ala Gly Ser Tyr Ala
85          90          95

```

```
Glu His Val Gly Leu Leu His Val Val Gly Val Pro Ser Ile Ser Ser
100         105         110

```

```
Gln Ala Lys Gln Leu Leu Leu His His Thr Leu Gly Asn Gly Asp Phe
Page 168

```

115

120

125

Thr Val Phe His Arg Met Ser Asn Asn Ile Ser Gln Thr Thr Ala Phe
 130 135 140
 Ile Ser Asp Ile Asn Ser Ala Pro Ala Glu Ile Asp Arg Cys Ile Arg
 145 150 155 160
 Glu Ala Tyr Val Lys Gln Arg Pro Val Tyr Ile Gly Leu Pro Ala Asn
 165 170 175
 Leu Val Asp Leu Asn Val Pro Ala Ser Leu Leu Glu Ser Pro Ile Asn
 180 185 190
 Leu Ser Leu Glu Lys Asn Asp Pro Glu Ala Gln Asp Glu Val Ile Asp
 195 200 205
 Ser Val Leu Asp Leu Ile Lys Lys Ser Ser Asn Pro Ile Ile Leu Val
 210 215 220
 Asp Ala Cys Ala Ser Arg His Asp Cys Lys Ala Glu Val Thr Gln Leu
 225 230 235 240
 Ile Glu Gln Thr Gln Phe Pro Val Phe Val Thr Pro Met Gly Lys Gly
 245 250 255
 Thr Val Asp Glu Gly Gly Val Asp Gly Glu Leu Leu Glu Asp Asp Pro
 260 265 270
 His Leu Ile Ala Lys Val Ala Ala Arg Leu Ser Ala Gly Lys Asn Ala
 275 280 285
 Ala Ser Arg Phe Gly Gly Val Tyr Val Gly Thr Leu Ser Lys Pro Glu
 290 295 300
 Val Lys Asp Ala Val Glu Ser Ala Asp Leu Ile Leu Ser Val Gly Ala
 305 310 315 320
 Leu Leu Ser Asp Phe Asn Thr Gly Ser Phe Ser Tyr Ser Tyr Arg Thr
 325 330 335
 Lys Asn Ile Val Glu Phe His Ser Asp Tyr Thr Lys Ile Arg Gln Ala
 340 345 350
 Thr Phe Pro Gly Val Gln Met Lys Glu Ala Leu Gln Glu Leu Asn Lys
 355 360 365
 Lys Val Ser Ser Ala Ala Ser His Tyr Glu Val Lys Pro Val Pro Lys
 370 375 380
 Ile Lys Leu Ala Asn Thr Pro Ala Thr Arg Glu Val Lys Leu Thr Gln

385 390 395 400

Glu Trp Leu Trp Thr Arg Val Ser Ser Trp Phe Arg Glu Gly Asp Ile
405 410 415

Ile Ile Thr Glu Thr Gly Thr Ser Ser Phe Gly Ile Val Gln Ser Arg
420 425 430

Phe Pro Asn Asn Thr Ile Gly Ile Ser Gln Val Leu Trp Gly Ser Ile
435 440 445

Gly Phe Ser Val Gly Ala Thr Leu Gly Ala Ala Met Ala Ala Gln Glu
450 455 460

Leu Asp Pro Asn Lys Arg Thr Ile Leu Phe Val Gly Asp Gly Ser Leu
465 470 475 480

Gln Leu Thr Val Gln Glu Ile Ser Thr Ile Ile Arg Trp Gly Thr Thr
485 490 495

Pro Tyr Leu Phe Val Leu Asn Asn Asp Gly Tyr Thr Ile Glu Arg Leu
500 505 510

Ile His Gly Val Asn Ala Ser Tyr Asn Asp Ile Gln Pro Trp Gln Asn
515 520 525

Leu Glu Ile Leu Pro Thr Phe Ser Ala Lys Asn Tyr Asp Ala Val Arg
530 535 540

Ile Ser Asn Ile Gly Glu Ala Glu Asp Ile Leu Lys Asp Lys Glu Phe
545 550 555 560

Gly Lys Asn Ser Lys Ile Arg Leu Ile Glu Val Met Leu Pro Arg Leu
565 570 575

Asp Ala Pro Ser Asn Leu Ala Lys Gln Ala Ala Ile Thr Ala Ala Thr
580 585 590

Asn Ala Glu Ala
595

<210> 118
<211> 1788
<212> DNA
<213> Pichia stipitis

<400> 118
atggctgaag tctcattagg aagatatctc ttcgagagat tgtaccaatt gcaagtgcag 60
accatcttcg gtgtccctgg tgatttcaac ttgtcgcttt tggacaagat ctacgaagtg 120
gaagatgccc atggcaagaa ttcgtttaga tgggctggta atgccaacga attgaatgca 180
tcgtacgctg ctgacggtta ctcgagagtc aagcgtttag ggtgtttggt cactaccttt 240

```

gggtgctggtg aattgtctgc tttgaatggt attgccggtt cttatgccga acatgttggt      300
ttgcttcatg tcgtaggtgt tccatcgatt tcctcgcaag ctaagcaatt gttacttcac      360
cacactttgg gtaatggtga tttcactggt ttccatagaa tgtccaacaa catttctcag      420
accacagcct ttatctccga tatcaactcg gctccagctg aaattgatag atgtatcaga      480
gaggcctacg tcaaacaag accagtttat atcgggttac cagctaactt agttgatttg      540
aatgttccgg cctctttgct tgagtctcca atcaacttgt cgttggaaaa gaacgaccca      600
gaggctcaag atgaagtcac tgactctgtc ttagacttga tcaaaaagtc gctgaaccca      660
atcatcttgg tcgatgcctg tgcctcgaga catgactgta aggctgaagt tactcagttg      720
attgaacaaa cccaattccc agtatttgtc actccaatgg gtaaaggtag cgttgatgag      780
ggtggtgtag acggagaatt gttagaagat gatcctcatt tgattgccaa ggtcgctgct      840
aggttgtctg ctggcaagaa cgctgcctct agattcggag gtgtttatgt cggaaccttg      900
tcgaagcccg aagtcaagga cgctgtagag agtgcagatt tgattttgtc tgtcggtgcc      960
cttttgtctg attcaacac tggttcattt tcctactcct acagaaccaa gaacatcgtc     1020
gaattccatt ctgattacac taagattaga caagccactt tcccagggtg gcagatgaag     1080
gaagccttgc aagaattgaa caagaaagt tcatctgctg ctagtcacta tgaagtcaag     1140
cctgtgcca agatcaagtt ggccaatata ccagccacca gagaagtcaa gttaactcag     1200
gaatgggtgt ggaccagagt gtcttcgtgg ttcagagaag gtgatattat tatcaccgaa     1260
accggtacat cctccttcgg tatagttcaa tccagattcc caaacaacac catcggtatc     1320
tccaagtat tgtgggggtt tattggtttc tctgttggtg ccactttggg tgctgccatg     1380
gctgccaag aactcgacct taacaagaga accatcttgt ttgttgaga tggttctttg     1440
caattgaccg ttcaggaaat ctccaccata atcagatggg gtaccacacc ttaccttttc     1500
gtgttgaaca atgacggtta caccatcgag cgtttgatcc acggtgtaaa tgcctcatat     1560
aatgacatcc aaccatggca aaacttgga atcttgcccta ctttctcggc caagaactac     1620
gacgctgtga gaatctcaa catcgagaa gcagaagata tcttgaaaga caaggaattc     1680
ggaaagaact ccaagattag attgatagaa gtcattgtac caagattgga tgcaccatct     1740
aaccttgcca aacaagctgc cattacagct gccaccaacg ccgaagct      1788

```

<210> 119
<211> 569
<212> PRT
<213> P i c h i a s t i p i t i s

<400> 119

Met Val Ser Thr Tyr Pro Glu Ser Glu Val Thr Leu Gly Arg Tyr Leu
1 5 10 15

Phe Glu Arg Leu His Gln Leu Lys Val Asp Thr Ile Phe Gly Leu Pro
20 25 30

20130927_CL5646W0PCT_ST25. txt

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

20130927_CL5646W0PCT_ST25. txt

Ser Tyr Lys Thr Lys Asn Val Val Glu Phe His Ser Asp Tyr Met Lys
305 310 315 320

Ile Arg Gln Ala Thr Phe Pro Gly Val Gln Met Lys Glu Ala Leu Gln
325 330 335

Gln Leu Ile Lys Arg Val Ser Ser Tyr Ile Asn Pro Ser Tyr Ile Pro
340 345 350

Thr Arg Val Pro Lys Arg Lys Gln Pro Leu Lys Ala Pro Ser Glu Ala
355 360 365

Pro Leu Thr Gln Glu Tyr Leu Trp Ser Lys Val Ser Gly Trp Phe Arg
370 375 380

Glu Gly Asp Ile Ile Val Thr Glu Thr Gly Thr Ser Ala Phe Gly Ile
385 390 395 400

Ile Gln Ser His Phe Pro Ser Asn Thr Ile Gly Ile Ser Gln Val Leu
405 410 415

Trp Gly Ser Ile Gly Phe Thr Val Gly Ala Thr Val Gly Ala Ala Met
420 425 430

Ala Ala Gln Glu Ile Asp Pro Ser Arg Arg Val Ile Leu Phe Val Gly
435 440 445

Asp Gly Ser Leu Gln Leu Thr Val Gln Glu Ile Ser Thr Leu Cys Lys
450 455 460

Trp Asp Cys Asn Asn Thr Tyr Leu Tyr Val Leu Asn Asn Asp Gly Tyr
465 470 475 480

Thr Ile Glu Arg Leu Ile His Gly Lys Ser Ala Ser Tyr Asn Asp Ile
485 490 495

Gln Pro Trp Asn His Leu Ser Leu Leu Arg Leu Phe Asn Ala Lys Lys
500 505 510

Tyr Gln Asn Val Arg Val Ser Thr Ala Gly Glu Leu Asp Ser Leu Phe
515 520 525

Ser Asp Lys Lys Phe Ala Ser Pro Asp Arg Ile Arg Met Ile Glu Val
530 535 540

Met Leu Ser Arg Leu Asp Ala Pro Ala Asn Leu Val Ala Gln Ala Lys
545 550 555 560

Leu Ser Glu Arg Val Asn Leu Glu Asn
565

<210> 120
 <211> 1707
 <212> DNA
 <213> *Pichia stipitis*

<400> 120
 atggtatcaa cctaccaga atcagagggtt actctaggaa ggtacctctt tgagcgactc 60
 caccaattga aagtggacac ctttttcggc ttgccgggtg acttcaacct ttccttattg 120
 gacaaagtgt atgaagttcc ggatatgagg tgggctggaa atgccaacga attgaatgct 180
 gcctatgctg ccgatgggta ctccagaata aagggattgt cttgcttggt cacaactttt 240
 ggtgttggtg aattgtctgc tttaaacgga gttggtggtg cctatgctga acacgtagga 300
 cttctacatg tcgttggagt tccatccata tcgtcacagg ctaaacagtt gttgctccac 360
 cataccttgg gtaatggtga cttactgtt tttcacagaa tgtccaatag catttctcaa 420
 actacagcat ttctctcaga tatctctatt gcaccaggctc aaatagatag atgcatcaga 480
 gaagcatatg ttcacagag accagtttat gttggtttac cggcaaatat ggttgatctc 540
 aaggttcctt ctagtctctt agaaactcca attgatttga aattgaaaca aaatgatcct 600
 gaagctcaag aagttgttga aacagtcctg aagttggtgt cccaagctac aaacccatt 660
 atcttggtag acgcttgtgc cctcagacac aattgcaaag aggaagtcaa acaattgggt 720
 gatgccacta atttcaagt ctttacaact ccaatgggta aatctgggtat ctccgaatct 780
 catccaagat tgggcgggtg ctatgtcggg acaatgtcga gtcctcaagt caaaaaagcc 840
 gttgaaaatg ccgatcttat actatctgtt ggttcgttgt tatcggactt caatacaggt 900
 tcattttcat actcctacaa gacgaagaat gttgttgaat tccactctga ctatatgaaa 960
 atcagacagg ccaccttccc aggagttcaa atgaaagaag ccttgcaaca gttgataaaa 1020
 agggctctctt cttacatcaa tccaagctac attcctactc gagttcctaa aaggaaacag 1080
 ccattgaaag ctccatcaga agtcctttg acccaagaat atttgtgggtc taaagtatcc 1140
 ggctgggtta gagagggtga tattatcgta accgaaactg gtacatctgc tttcggaatt 1200
 attcaatccc attttcccag caacactatc ggtatatccc aagtcttgtg gggctcaatt 1260
 ggtttcacag taggtgcaac agttggtgct gccatggcag cccaggaaat cgaccctagc 1320
 aggagagtaa ttttgttcgt cggtgatggt tcattgcagt tgacggttca ggaaatctct 1380
 acgttgtgta aatgggattg taacaatact tatctttacg tgttgaacaa tgatgggttac 1440
 actatagaaa ggttgatcca cggcaaaagt gccagctaca acgatataca gccttggaac 1500
 catttatcct tgcttcgctt attcaatgct aagaaatacc aaaatgtcag agtatcgact 1560
 gctggagaat tggactcttt gttctctgat aagaaatttg cttctccaga taggataaga 1620
 atgattgagg tgatgttatc gagattggat gcaccagcaa atcttgttgc tcaagcaaag 1680
 ttgtctgaac gggtaaacct tgaaaat 1707

<210> 121
 <211> 563

<212> PRT

<213> Kluyveromyces fragilis

<400> 121

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

20130927_CL5646WOPCT_ST25. txt

Pro Arg Phe Gly Gly Val Tyr Val Gly Thr Leu Ser Ser Pro Ala Val
260 265 270

Lys Glu Ala Val Glu Ser Ala Asp Leu Val Leu Ser Val Gly Ala Leu
275 280 285

Leu Ser Asp Phe Asn Thr Gly Ser Phe Ser Tyr Ser Tyr Lys Thr Lys
290 295 300

Asn Ile Val Glu Phe His Ser Asp Tyr Thr Lys Ile Arg Ser Ala Thr
305 310 315 320

Phe Pro Gly Val Gln Met Lys Phe Ala Leu Gln Lys Leu Leu Thr Lys
325 330 335

Val Ala Asp Ala Ala Lys Gly Tyr Lys Pro Val Pro Val Pro Ser Glu
340 345 350

Pro Glu His Asn Glu Ala Val Ala Asp Ser Thr Pro Leu Lys Gln Glu
355 360 365

Trp Val Trp Thr Gln Val Gly Glu Phe Leu Arg Glu Gly Asp Val Val
370 375 380

Ile Thr Glu Thr Gly Thr Ser Ala Phe Gly Ile Asn Gln Thr His Phe
385 390 395 400

Pro Asn Asn Thr Tyr Gly Ile Ser Gln Val Leu Trp Gly Ser Ile Gly
405 410 415

Phe Thr Thr Gly Ala Thr Leu Gly Ala Ala Phe Ala Ala Glu Glu Ile
420 425 430

Asp Pro Lys Lys Arg Val Ile Leu Phe Ile Gly Asp Gly Ser Leu Gln
435 440 445

Leu Thr Val Gln Glu Ile Ser Thr Met Ile Arg Trp Gly Leu Lys Pro
450 455 460

Tyr Leu Phe Val Leu Asn Asn Asp Gly Tyr Thr Ile Glu Arg Leu Ile
465 470 475 480

His Gly Glu Thr Ala Gln Tyr Asn Cys Ile Gln Asn Trp Gln His Leu
485 490 495

Glu Leu Leu Pro Thr Phe Gly Ala Lys Asp Tyr Glu Ala Val Arg Val
500 505 510

Ser Thr Thr Gly Glu Trp Asn Lys Leu Thr Thr Asp Glu Lys Phe Gln
515 520 525

Asp Asn Thr Arg Ile Arg Leu Ile Glu Val Met Leu Pro Thr Met Asp
 530 535 540

Ala Pro Ser Asn Leu Val Lys Gln Ala Gln Leu Thr Ala Ala Thr Asn
 545 550 555 560

Ala Lys Asn

<210> 122
 <211> 1689
 <212> DNA
 <213> Kluyveromyces lactic

<400> 122
 atgtctgaaa ttacattagg tcgttacttg ttcgaaagat taaagcaagt cgaagttcaa 60
 accatctttg gtctaccagg tgatttcaac ttgtccctat tggacaatat ctacgaagtc 120
 ccaggatatga gatgggctgg taatgccaac gaattgaacg ctgcttacgc tgctgatggt 180
 tacgccagat taaagggat gtcctgtatc atcaccacct tcggtgtcgg tgaattgtct 240
 gctttgaacg gtattgccgg ttcttacgct gaacacgttg gtgtcttgca cgttgtcgg 300
 gttccatccg tctcttctca agctaagcaa ttgttgttgc accacacctt gggtaacgg 360
 gacttactg tttccacag aatgtcctcc aacatttctg aaaccactgc tatgatcacc 420
 gatatcaaca ctgccccagc tgaaatcgac agatgtatca gaaccactta cgtttcccaa 480
 agaccagtct acttgggttt gccagctaac ttgggtcgact tgactgtccc agcttctttg 540
 ttggacactc caattgattt gagcttgaag ccaaagacc cagaagccga agaagaagtc 600
 atcgaaaacg tcttgcaact gatcaaggaa gctaagaacc cagttatctt ggctgatgct 660
 tgttgttcca gacacgatgc caaggctgag accaagaagt tgatcgactt gactcaattc 720
 ccagccttcg ttaccccaat gggtaagggt tccattgacg aaaagcacc aagattcgg 780
 ggtgtctacg tcggtaccct atcttctcca gctgtcaagg aagccgttga atctgctgac 840
 ttggttctat cggtcggtgc tctattgtcc gatttcaaca ctggttcttt ctcttactct 900
 tacaagacca agaacattgt cgaattccac tctgactaca ccaagatcag aagcgctacc 960
 ttcccagggtg tccaaatgaa gttcgcttta caaaaattgt tgactaaggt tgccgatgct 1020
 gctaaggggtt acaagccagt tccagttcca tctgaaccag aacacaacga agctgtcgt 1080
 gactccactc cattgaagca agaatgggtc tggactcaag tcggtgaatt cttgagagaa 1140
 ggtgatgttg ttatcactga aaccggtacc tctgccttcg gtatcaacca aactcatttc 1200
 ccaaacaaca catacggat ctctcaagtt ttatgggggt ccattgggtt caccactggt 1260
 gctaccttg gtgctgcctt cgctgccgaa gaaattgatc caaagaagag agttatctta 1320
 ttcatgggtg acggttcttt gcaattgact gttcaagaaa tctccaccat gatcagatgg 1380
 ggcttgaagc catacttggt cgtattgaac aacgacgggt acaccattga aagattgatt 1440
 cacgggtgaaa ccgctcaata caactgtatc caaaactggc aacacttgga attattgcca 1500

actttcgggtg ccaaggacta cgaagctgtc agagtttcca ccactggtga atggaacaag 1560
 ttgaccactg acgaaaagtt ccaagacaac accagaatca gattgatcga agttatgttg 1620
 ccaactatgg atgctccatc taacttggtt aagcaagctc aattgactgc tgctaccaac 1680
 gctaagaac 1689

<210> 123
 <211> 571
 <212> PRT
 <213> Yarrowia lipolytica

<400> 123

Met Ser Asp Ser Glu Pro Gln Met Val Asp Leu Gly Asp Tyr Leu Phe
 1 5 10 15

Ala Arg Phe Lys Gln Leu Gly Val Asp Ser Val Phe Gly Val Pro Gly
 20 25 30

Asp Phe Asn Leu Thr Leu Leu Asp His Val Tyr Asn Val Asp Met Arg
 35 40 45

Trp Val Gly Asn Thr Asn Glu Leu Asn Ala Gly Tyr Ser Ala Asp Gly
 50 55 60

Tyr Ser Arg Val Lys Arg Leu Ala Cys Leu Val Thr Thr Phe Gly Val
 65 70 75 80

Gly Glu Leu Ser Ala Val Ala Ala Val Ala Gly Ser Tyr Ala Glu His
 85 90 95

Val Gly Val Val His Val Val Gly Val Pro Ser Thr Ser Ala Glu Asn
 100 105 110

Lys His Leu Leu Leu His His Thr Leu Gly Asn Gly Asp Phe Arg Val
 115 120 125

Phe Ala Gln Met Ser Lys Leu Ile Ser Glu Tyr Thr His His Ile Glu
 130 135 140

Asp Pro Ser Glu Ala Ala Asp Val Ile Asp Thr Ala Ile Arg Ile Ala
 145 150 155 160

Tyr Thr His Gln Arg Pro Val Tyr Ile Ala Val Pro Ser Asn Phe Ser
 165 170 175

Glu Val Asp Ile Ala Asp Gln Ala Arg Leu Asp Thr Pro Leu Asp Leu
 180 185 190

Ser Leu Gln Pro Asn Asp Pro Glu Ser Gln Tyr Glu Val Ile Glu Glu
 195 200 205

Ile Cys Ser Arg Ile Lys Ala Ala Lys Lys Pro Val Ile Leu Val Asp
 210 215 220
 Ala Cys Ala Ser Arg Tyr Arg Cys Val Asp Glu Thr Lys Glu Leu Ala
 225 230 235 240
 Lys Ile Thr Asn Phe Ala Tyr Phe Val Thr Pro Met Gly Lys Gly Ser
 245 250 255
 Val Asp Glu Asp Thr Asp Arg Tyr Gly Gly Thr Tyr Val Gly Ser Leu
 260 265 270
 Thr Ala Pro Ala Thr Ala Glu Val Val Glu Thr Ala Asp Leu Ile Ile
 275 280 285
 Ser Val Gly Ala Leu Leu Ser Asp Phe Asn Thr Gly Ser Phe Ser Tyr
 290 295 300
 Ser Tyr Ser Thr Lys Asn Val Val Glu Leu His Ser Asp His Val Lys
 305 310 315 320
 Ile Lys Ser Ala Thr Tyr Asn Asn Val Gly Met Lys Met Leu Phe Pro
 325 330 335
 Pro Leu Leu Glu Ala Val Lys Lys Leu Val Ala Glu Thr Pro Asp Phe
 340 345 350
 Ala Ser Lys Ala Leu Ala Val Pro Asp Thr Thr Pro Lys Ile Pro Glu
 355 360 365
 Val Pro Asp Asp His Ile Thr Thr Gln Ala Trp Leu Trp Gln Arg Leu
 370 375 380
 Ser Tyr Phe Leu Arg Pro Thr Asp Ile Val Val Thr Glu Thr Gly Thr
 385 390 395 400
 Ser Ser Phe Gly Ile Ile Gln Thr Lys Phe Pro His Asn Val Arg Gly
 405 410 415
 Ile Ser Gln Val Leu Trp Gly Ser Ile Gly Tyr Ser Val Gly Ala Ala
 420 425 430
 Cys Gly Ala Ser Ile Ala Ala Gln Glu Ile Asp Pro Gln Gln Arg Val
 435 440 445
 Ile Leu Phe Val Gly Asp Gly Ser Leu Gln Leu Thr Val Thr Glu Ile
 450 455 460
 Ser Cys Met Ile Arg Asn Asn Val Lys Pro Tyr Ile Phe Val Leu Asn
 465 470 475 480

20130927_CL5646WOPCT_ST25.txt

Asn Asp Gly Tyr Thr Ile Glu Arg Leu Ile His Gly Glu Asn Ala Ser
485 490 495

Tyr Asn Asp Val His Met Trp Lys Tyr Ser Lys Ile Leu Asp Thr Phe
500 505 510

Asn Ala Lys Ala His Glu Ser Ile Val Val Asn Thr Lys Gly Glu Met
515 520 525

Asp Ala Leu Phe Asp Asn Glu Glu Phe Ala Lys Pro Asp Lys Ile Arg
530 535 540

Leu Ile Glu Val Met Cys Asp Lys Met Asp Ala Pro Ala Ser Leu Ile
545 550 555 560

Lys Gln Ala Glu Leu Ser Ala Lys Thr Asn Val
565 570

<210> 124
<211> 1713
<212> DNA
<213> *Yarrowia lipolytica*

<400> 124
atgagcgact ccgaaccca aatggtcgac ctgggcgact atctctttgc ccgattcaag 60
cagctaggcg tggactccgt ctttgagtg cccggcgact tcaacctcac cctgttggac 120
cacgtgtaca atgtcgacat gcggtgggtt gggaacacaa acgagctgaa tgccggctac 180
tcggccgacg gctactcccg ggtcaagcgg ctggcatgtc ttgtcaccac ctttggcgtg 240
ggagagctgt ctgccgtggc tgctgtggca ggctcgtacg ccgagcatgt gggcgtgggtg 300
catgttgtgg gcgttcccag cacctctgct gagaacaagc atctgctgct gcaccacaca 360
ctcggtaacg gcgacttccg ggtctttgcc cagatgtcca aactcatctc cgagtacacc 420
caccatattg aggaccccag cgaggctgcc gacgtaatcg acaccgccat ccgaatcgcc 480
tacaccacc agcggcccggt ttacattgct gtgccctcca acttctccga ggtcgatatt 540
gccgaccagg ctagactgga tacccccctg gacctttcgc tgcagcccaa cgaccccgag 600
agccagtacg aggtgattga ggagatttgc tcgcgtatca aggccgcaa gaagcccgtg 660
attctcgtcg acgcctgcgc ttcgcgatac agatgtgtgg acgagaccaa ggagctggcc 720
aagatcacca actttgccta ctttgtcact ccatgggta agggttctgt ggacgaggat 780
actgaccggt acggaggaa atacgtcgga tcgctgactg ctctgtctac tgccgaggtg 840
gttgagacag ctgatctcat catctccgta ggagctcttc tgtcggactt caacaccggt 900
tccttctcgt actctactc caccaaaaac gtgggtggaat tgcattcgga ccacgtcaaa 960
atcaagtccg ccacctacaa caacgtcggc atgaaaatgc tgttcccgcc cctgctcgaa 1020
gccgtcaaga aactggttgc cgagaccct gactttgcat ccaaggctct ggctgttccc 1080
gacaccactc ccaagatccc cgaggtagcc gatgatcaca ttacgacca ggcattggctg 1140

20130927_CL5646WOPCT_ST25.txt

ttgcagcgtc tcagttactt tctgaggccc accgacatcg tggtcaccga gaccggaacc	1200
tcgtccctttg gaatcatcca gaccaagttc cccacaacg tccgaggtat ctgcgaggtg	1260
ctgtgggggct ctattggata ctcggtggga gcagcctgtg gagcctccat tgctgcacag	1320
gagattgacc cccagcagcg agtgattctg tttgtgggcg acggctctct tcagctgacg	1380
gtgaccgaga tctcgtgcat gatccgcaac aacgtcaagc cgtacatttt tgtgctcaac	1440
aacgacggct acaccatcga gaggtcatt cacggcgaaa acgcctcgta caacgatgtg	1500
cacatgtgga agtactccaa gattctcgac acgttcaacg ccaaggccca cgagtcgatt	1560
gtgggtcaaca ccaagggcga gatggacgct ctgttcgaca acgaagagtt tgccaagccc	1620
gacaagatcc ggctcattga ggtcatgtgc gacaagatgg acgcgcctgc ctcgttgatc	1680
aagcaggctg agctctctgc caagaccaac gtt	1713

<210> 125
 <211> 571
 <212> PRT
 <213> Schizosaccharomyces pombe

<400> 125

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

Gln	Leu	Gly	Val	Lys	Ser	Ile	Leu	Gly	Val	Pro	Gly	Asp	Phe	Asn	Leu
			20					25					30		

Ala	Leu	Leu	Asp	Leu	Ile	Glu	Lys	Val	Gly	Asp	Glu	Lys	Phe	Arg	Trp
		35					40					45			

Val	Gly	Asn	Thr	Asn	Glu	Leu	Asn	Gly	Ala	Tyr	Ala	Ala	Asp	Gly	Tyr
	50					55					60				

Ala	Arg	Val	Asn	Gly	Leu	Ser	Ala	Ile	Val	Thr	Thr	Phe	Gly	Val	Gly
65					70					75					80

Glu	Leu	Ser	Ala	Ile	Asn	Gly	Val	Ala	Gly	Ser	Tyr	Ala	Glu	His	Val
				85					90					95	

Pro	Val	Val	His	Ile	Val	Gly	Met	Pro	Ser	Thr	Lys	Val	Gln	Asp	Thr
			100					105					110		

Gly	Ala	Leu	Leu	His	His	Thr	Leu	Gly	Asp	Gly	Asp	Phe	Arg	Thr	Phe
		115					120					125			

Met	Asp	Met	Phe	Lys	Lys	Val	Ser	Ala	Tyr	Ser	Ile	Met	Ile	Asp	Asn
	130					135					140				

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

Lys Lys Ala Arg Pro Val Tyr Ile Gly Ile Pro Ser Asp Ala Gly Tyr
 165 170 175
 Phe Lys Ala Ser Ser Ser Asn Leu Gly Lys Arg Leu Lys Leu Glu Glu
 180 185 190
 Asp Thr Asn Asp Pro Ala Val Glu Gln Glu Val Ile Asn His Ile Ser
 195 200 205
 Glu Met Val Val Asn Ala Lys Lys Pro Val Ile Leu Ile Asp Ala Cys
 210 215 220
 Ala Val Arg His Arg Val Val Pro Glu Val His Glu Leu Ile Lys Leu
 225 230 235 240
 Thr His Phe Pro Thr Tyr Val Thr Pro Met Gly Lys Ser Ala Ile Asp
 245 250 255
 Glu Thr Ser Gln Phe Phe Asp Gly Val Tyr Val Gly Ser Ile Ser Asp
 260 265 270
 Pro Glu Val Lys Asp Arg Ile Glu Ser Thr Asp Leu Leu Leu Ser Ile
 275 280 285
 Gly Ala Leu Lys Ser Asp Phe Asn Thr Gly Ser Phe Ser Tyr His Leu
 290 295 300
 Ser Gln Lys Asn Ala Val Glu Phe His Ser Asp His Met Arg Ile Arg
 305 310 315 320
 Tyr Ala Leu Tyr Pro Asn Val Ala Met Lys Tyr Ile Leu Arg Lys Leu
 325 330 335
 Leu Lys Val Leu Asp Ala Ser Met Cys His Ser Lys Ala Ala Pro Thr
 340 345 350
 Ile Gly Tyr Asn Ile Lys Pro Lys His Ala Glu Gly Tyr Ser Ser Asn
 355 360 365
 Glu Ile Thr His Cys Trp Phe Trp Pro Lys Phe Ser Glu Phe Leu Lys
 370 375 380
 Pro Arg Asp Val Leu Ile Thr Glu Thr Gly Thr Ala Asn Phe Gly Val
 385 390 395 400
 Leu Asp Cys Arg Phe Pro Lys Asp Val Thr Ala Ile Ser Gln Val Leu
 405 410 415
 Trp Gly Ser Ile Gly Tyr Ser Val Gly Ala Met Phe Gly Ala Val Leu
 420 425 430

20130927_CL5646WOPCT_ST25. txt

Ala Val His Asp Ser Lys Glu Pro Asp Arg Arg Thr Ile Leu Val Val
435 440 445

Gly Asp Gly Ser Leu Gln Leu Thr Ile Thr Glu Ile Ser Thr Cys Ile
450 455 460

Arg His Asn Leu Lys Pro Ile Ile Phe Ile Ile Asn Asn Asp Gly Tyr
465 470 475 480

Thr Ile Glu Arg Leu Ile His Gly Leu His Ala Ser Tyr Asn Glu Ile
485 490 495

Asn Thr Lys Trp Gly Tyr Gln Gln Ile Pro Lys Phe Phe Gly Ala Ala
500 505 510

Glu Asn His Phe Arg Thr Tyr Cys Val Lys Thr Pro Thr Asp Val Glu
515 520 525

Lys Leu Phe Ser Asp Lys Glu Phe Ala Asn Ala Asp Val Ile Gln Val
530 535 540

Val Glu Leu Val Met Pro Met Leu Asp Ala Pro Arg Val Leu Val Glu
545 550 555 560

Gln Ala Lys Leu Thr Ser Lys Ile Asn Lys Gln
565 570

<210> 126
<211> 1713
<212> DNA
<213> Schizosaccharomyces pombe

<400> 126
atgagtggg atatttagt cggatgaatat ctattcaaaa ggcttgaaca attaggggtc 60
aagtcattc ttggtgttc aggagatttc aatttagctc tacttgactt aattgagaaa 120
gttgagatg agaaatttcg ttgggttggc aataccaatg agttgaatgg tgcttatgcc 180
gctgatggtt atgctcgtgt taatggtctt tcagccattg ttacaacgtt cggcgtggga 240
gagctttccg ctattaatgg agtggcaggt tcttatgcgg agcatgtccc agtagttcat 300
attgttgga tgccttcac aaaggtgcaa gatactggag ctttgcttca tcatacttta 360
ggagatggag actttcgac tttcatggat atgtttaaga aagtttctgc ctacagtata 420
atgatcgata acggaacga tgcagctgaa aagatcgatg aagccttgct gatttgttat 480
aaaaaggcta ggcctgttta cattggtatt ccttctgatg ctggctactt caaagcatct 540
tcatcaaatc ttgggaaaag actaaagctc gaggaggata ctaacgatcc agcagttgag 600
caagaagtca tcaatcatat ctcggaaatg gttgtcaatg caaagaaacc agtgatttta 660
attgacgctt gtgctgtaag acatcgtgtc gttccagaag tacatgagct gattaaattg 720
acccatttcc ctacatatgt aactcccatg ggtaaactctg caattgacga aacttcgcaa 780

```

ttttttgacg gcgtttatgt tggttcaatt tcagatcctg aagttaaaga cagaattgaa      840
tccactgatc tgttgctatc catcggtgct ctcaaacag actttaacac gggttccttc      900
tcttaccacc tcagccaaaa gaatgccgtt gagtttcatt cagaccacat gcgcattcga      960
tatgctcttt atccaaatgt agccatgaag tatattcttc gcaaactggt gaaagtactt    1020
gatgcttcta tgtgtcattc caaggctgct cctaccattg gctacaacat caagcctaag    1080
catgcggaag gatattcttc caacgagatt actcattgct ggttttggcc taaatttagt    1140
gaatttttga agccccgaga tgttttgatc accgagactg gaactgcaaa ctttggtgtc    1200
cttgattgca ggtttccaaa ggatgtaaca gccatttccc aggtattatg gggatctatt    1260
ggataactccg ttggtgcaat gtttggtgct gttttggccg tccacgattc taaagagccc    1320
gatcgtcgta ccattcttgt agtaggtgat ggatccttac aactgacgat tacagagatt    1380
tcaacctgca ttcgccataa cctcaaacca attattttca taattaacaa cgacggttac    1440
accattgagc gtttaattca tggtttgcatt gctagctata acgaaattaa cactaaatgg    1500
ggctaccaac agattcccaa gtttttcgga gctgctgaaa accacttccg cacttactgt    1560
gttaaaaactc ctactgacgt tgaaaagttg tttagcgaca aggagtttgc aaatgcagat    1620
gtcattcaag tagttgagct tgtaatgcct atgttggatg cacctcgtgt cctagttgag    1680
caagccaagt tgacgtctaa gatcaataag caa                                1713

```

```

<210> 127
<211> 563
<212> PRT
<213> Zygosaccharomyces rouxi i

```

```
<400> 127
```

```
Met Ser Glu Ile Thr Leu Gly Arg Tyr Leu Phe Glu Arg Leu Lys Gln
1           5           10          15
```

```
Val Asp Thr Asn Thr Ile Phe Gly Val Pro Gly Asp Phe Asn Leu Ser
20           25           30
```

```
Leu Leu Asp Lys Val Tyr Glu Val Gln Gly Leu Arg Trp Ala Gly Asn
35           40           45
```

```
Ala Asn Glu Leu Asn Ala Ala Tyr Ala Ala Asp Gly Tyr Ala Arg Val
50           55           60
```

```
Lys Gly Leu Ala Ala Leu Ile Thr Thr Phe Gly Val Gly Glu Leu Ser
65           70           75           80
```

```
Ala Leu Asn Gly Ile Ala Gly Ser Tyr Ala Glu His Val Gly Val Leu
85           90           95
```

```
His Ile Val Gly Val Pro Ser Val Ser Ser Gln Ala Lys Gln Leu Leu
100          105          110
```

20130927_CL5646WOPCT_ST25. txt

Leu His His Thr Leu Gly Asn Gly Asp Phe Thr Val Phe His Arg Met
 115 120 125
 Ser Ala Asn Ile Ser Glu Thr Thr Ala Met Leu Thr Asp Ile Thr Ala
 130 135 140
 Ala Pro Ala Glu Ile Asp Arg Cys Ile Arg Val Ala Tyr Val Asn Gln
 145 150 155 160
 Arg Pro Val Tyr Leu Gly Leu Pro Ala Asn Leu Val Asp Gln Lys Val
 165 170 175
 Pro Ala Ser Leu Leu Asn Thr Pro Ile Asp Leu Ser Leu Lys Glu Asn
 180 185 190
 Asp Pro Glu Ala Glu Thr Glu Val Val Asp Thr Val Leu Glu Leu Ile
 195 200 205
 Lys Glu Ala Lys Asn Pro Val Ile Leu Ala Asp Ala Cys Cys Ser Arg
 210 215 220
 His Asp Val Lys Ala Glu Thr Lys Lys Leu Ile Asp Leu Thr Gln Phe
 225 230 235 240
 Pro Ser Phe Val Thr Pro Met Gly Lys Gly Ser Ile Asp Glu Gln Asn
 245 250 255
 Pro Arg Phe Gly Gly Val Tyr Val Gly Thr Leu Ser Ser Pro Glu Val
 260 265 270
 Lys Glu Ala Val Glu Ser Ala Asp Leu Val Leu Ser Val Gly Ala Leu
 275 280 285
 Leu Ser Asp Phe Asn Thr Gly Ser Phe Ser Tyr Ser Tyr Lys Thr Lys
 290 295 300
 Asn Val Val Glu Phe His Ser Asp His Ile Lys Ile Arg Asn Ala Thr
 305 310 315 320
 Phe Pro Gly Val Gln Met Lys Phe Val Leu Lys Lys Leu Leu Gln Ala
 325 330 335
 Val Pro Glu Ala Val Lys Asn Tyr Lys Pro Gly Pro Val Pro Ala Pro
 340 345 350
 Pro Ser Pro Asn Ala Glu Val Ala Asp Ser Thr Thr Leu Lys Gln Glu
 355 360 365
 Trp Leu Trp Arg Gln Val Gly Ser Phe Leu Arg Glu Gly Asp Val Val
 370 375 380

I l e Thr Gl u Thr Gly Thr Ser Al a Phe Gly I l e Asn Gl n Thr Hi s Phe
385 390 395 400

Pro Asn Gl n Thr Tyr Gly I l e Ser Gl n Val Leu Trp Gly Ser I l e Gly
405 410 415

Tyr Thr Thr Gly Ser Thr Leu Gly Al a Al a Phe Al a Al a Gl u Gl u I l e
420 425 430

Asp Pro Lys Lys Arg Val I l e Leu Phe I l e Gly Asp Gly Ser Leu Gl n
435 440 445

Leu Thr Val Gl n Gl u I l e Ser Thr Met I l e Arg Trp Gly Leu Lys Pro
450 455 460

Tyr Leu Phe Val Leu Asn Asn Asp Gly Tyr Thr I l e Gl u Arg Leu I l e
465 470 475 480

Hi s Gly Gl u Thr Al a Gl u Tyr Asn Cys I l e Gl n Pro Trp Lys Hi s Leu
485 490 495

Gl u Leu Leu Asn Thr Phe Gly Al a Lys Asp Tyr Gl u Asn Hi s Arg Val
500 505 510

Ser Thr Val Gly Gl u Trp Asn Lys Leu Thr Gl n Asp Pro Lys Phe Asn
515 520 525

Gl u Asn Ser Arg I l e Arg Met I l e Gl u Val Met Leu Gl u Val Met Asp
530 535 540

Al a Pro Ser Ser Leu Val Al a Gl n Al a Gl n Leu Thr Al a Al a Thr Asn
545 550 555 560

Al a Lys Gl n

<210> 128
<211> 1689
<212> DNA
<213> Zygosaccharomyces rouxi i

<400> 128
atgtctgaaa ttactctagg tcgttacttg ttcgaaagat taaagcaagt tgacactaac 60
accatcttcg gtgttcagg tgacttcaac ttgtccttgt tggacaaggt ctacgaagtg 120
caaggcttaa gatgggctgg taacgctaac gaattgaacg ctgcctacgc tgctgacggt 180
tacgccagag ttaagggttt ggctgctttg atcaccacct tcggtgtcgg tgaattgtct 240
gctttgaacg gtattgcagg ttcttacgct gaacacgttg gtgttttgca cattgttggt 300
gttccatctg tctcttctca agctaagcaa ttgttgttgc accacacctt gggtaacggt 360
gacttcactg ttttcacag aatgtccgcc aacatctctg aaaccaccgc tatgttgacc 420

20130927_CL5646WOPCT_ST25. txt

gacatcactg	ctgctccagc	tgaattgac	cgttgcatca	gagttgctta	cgtcaaccaa	480
agaccagtct	acttgggtct	accagctaac	ttggttgacc	aaaaggtccc	agcttctttg	540
ttgaacactc	caattgatct	atctctaaag	gagaacgacc	cagaagctga	aaccgaagtt	600
gttgacaccg	ttttggaatt	gatcaaggaa	gctaagaacc	cagttatctt	ggctgatgct	660
tgctgctcca	gacacgacgt	caaggctgaa	accaagaagt	tgatcgactt	gactcaattc	720
ccatctttcg	ttactcctat	gggtaagggt	tccatcgacg	aacaaaaccc	aagattcggg	780
gggtgtctacg	tcgggtactct	atccagccca	gaagttaagg	aagctgttga	atctgctgac	840
ttggttctat	ctgtcgggtg	tctattgtcc	gatttcaaca	ctggttcttt	ctcttactct	900
tacaagacca	agaacgttgt	tgaattccac	tctgaccaca	tcaagatcag	aaacgctacc	960
ttcccagggtg	ttcaaatgaa	attcgttttg	aagaaactat	tgcaagctgt	cccagaagct	1020
gtcaagaact	acaagccagg	tccagtccca	gctccgcat	ctccaaacgc	tgaagttgct	1080
gactctacca	ccttgaagca	agaatgggta	tggagacaag	tcggtagctt	cttgagagaa	1140
ggtagatgtg	ttattaccga	aactgggtacc	tctgctttcg	gtatcaacca	aactcacttc	1200
cctaaccaaa	cttacggtat	ctctcaagtc	ttgtgggggt	ctattgggta	caccactggg	1260
tccactttgg	gtgctgcctt	cgctgctgaa	gaaattgacc	ctaagaagag	agttatcttg	1320
ttcattgggtg	acggttctct	acaattgacc	gttcaagaaa	tctccaccat	gatcagatgg	1380
gggtctaaagc	catacttggt	cgttttgaac	aacgatgggt	acaccattga	aagattgatt	1440
cacggtgaaa	ccgctgaata	caactgtatc	caaccatgga	agcacttgga	attggttgaa	1500
accttcggtg	ccaaggacta	cgaaaaccac	agagtctcca	ctgtcgggtg	atggaacaag	1560
ttgactcaag	atccaaaatt	caacgaaaac	tctagaatta	gaatgatcga	agttatgctt	1620
gaagtcatgg	acgtccatc	ttctttggtc	gctcaagctc	aattgaccgc	tgctactaac	1680
gctaagcaa						1689

<210> 129
 <211> 4582
 <212> DNA
 <213> Arti f i c i a l Sequence

<220>
 <223> pLA54 pl asmi d

<400>	129		
gggtaccgag	ctcgaattca	ctggccgtcg	
ttttacaacg	tcgtgactgg	gaaaaccctg	60
gcgttaccca	acttaatcgc	cttgacgac	
atcccccttt	cgccagctgg	cgtaatagcg	120
aagaggcccc	caccgatcgc	ccttcccaac	
agttgcgcag	cctgaatggc	gaatggcgcc	180
tgatgcggta	ttttctcctt	acgcatctgt	
gcgggtatttc	acaccgcata	tggtgcactc	240
tcagtacaat	ctgctctgat	gccgcatagt	
taagccagcc	ccgacacccg	ccaacacccg	300
ctgacgcgcc	ctgacgggct	tgtctgtctc	
cggcatccgc	ttacagacaa	gctgtgaccg	360
tctccggggag	ctgcatgtgt	cagagggtttt	
caccgtcatc	accgaaacgc	gcgagacgaa	420

20130927_CL5646WOPCT_ST25. txt

agggcctcgt	gatacgccta	tttttatagg	ttaatgtcat	gataataatg	gtttcttaga	480
cgtcaggtgg	cacttttcgg	ggaaatgtgc	gcggaacccc	tatttgttta	tttttctaaa	540
tacattcaaa	tatgtatccg	ctcatgagac	aataaccctg	ataaatgctt	caataatatt	600
gaaaaaggaa	gagtatgagt	attcaacatt	tccgtgtcgc	ccttattccc	ttttttgcgg	660
cattttgcct	tcctgttttt	gctcacccag	aaacgctggt	gaaagtaaaa	gatgctgaag	720
atcagttggg	tgcacgagtg	ggttacatcg	aactggatct	caacagcggg	aagatccttg	780
agagttttcg	ccccgaagaa	cgttttccaa	tgatgagcac	ttttaagtt	ctgctatgtg	840
gcgcggtatt	atcccgtatt	gacgccgggc	aagagcaact	cggtcgccgc	atacactatt	900
ctcagaatga	cttggttag	tactcaccag	tcacagaaaa	gcatcttacg	gatggcatga	960
cagtaagaga	attatgcagt	gctgccataa	ccatgagtga	taacactgcg	gccaaacttac	1020
ttctgacaac	gatcggagga	ccgaaggagc	taaccgcttt	tttgcacaa	atgggggatac	1080
atgtaactcg	ccttgatcgt	tgggaaccgg	agctgaatga	gccataccaa	acgacgagcg	1140
tgacaccacg	atgcctgtag	caatggcaac	aacgttgccg	aaactattac	tggcgaacta	1200
cttactctag	cttcccggca	acaattaata	gactggatgg	aggcggataa	agttgcagga	1260
ccacttctgc	gctcggccct	tccggctggc	tggtttattg	ctgataaatc	tggagccggt	1320
gagcgtggtc	tcgcggtatc	attgcagcac	tggggccaga	tggtaagccc	tcccgtatcg	1380
tagttatcta	cacgacggga	gtcaggcaac	tatggatgaa	cgaaatagac	agatcgtgta	1440
gataggtgcc	tactgatta	agcattggta	actgtcagac	caagtttact	catatatact	1500
ttagattgat	ttaaaacttc	atttttaatt	taaaaggatc	taggtgaaga	tcctttttga	1560
taatctcatg	acaaaaatcc	cttaacgtga	gttttcgttc	cactgagcgt	cagaccccg	1620
agaaaagatc	aaaggatctt	cttgagatcc	tttttttctg	cgcgtaatct	gctgcttgca	1680
aacaaaaaaa	ccaccgctac	cagcggtggt	ttgtttgccg	gatcaagagc	taccaactct	1740
ttttccgaag	gtaactggct	tcagcagagc	gcagatacca	aatactgtcc	ttctagtgtg	1800
gccgtagtta	ggccaccact	tcaagaactc	tgtagcaccg	cctacatacc	tcgctctgct	1860
aatcctgtta	ccagtggctg	ctgccagtgg	cgataagtcg	tgtcttaccg	ggttggaactc	1920
aagacgatag	ttaccggata	aggcgcagcg	gtcgggctga	acgggggggt	cgtgcacaca	1980
gccagcttg	gagcgaacga	cctacaccga	actgagatac	ctacagcgtg	agctatgaga	2040
aagcgccacg	cttcccgaag	ggagaaaggc	ggacaggtat	ccggtaagcg	gcagggtcgg	2100
aacaggagag	cgcacgaggg	agcttcagg	gggaaacgcc	tggatatctt	atagtcctgt	2160
cgggtttcgc	cacctctgac	ttgagcgtcg	atttttgtga	tgctcgtcag	gggggcggag	2220
cctatggaaa	aacgccagca	acgcggcctt	tttacggttc	ctggcctttt	gctggccttt	2280
tgctcacatg	ttctttcctg	cgttatcccc	tgattctgtg	gataaccgta	ttaccgcctt	2340
tgagtgagct	gataccgctc	gccgcagccg	aacgaccgag	cgcagcgagt	cagtgagcga	2400
ggaagcggaa	gagcgcccaa	tacgcaaacc	gcctctcccc	gcgcgttggc	cgattcatta	2460

20130927_CL5646WOPCT_ST25. txt

atgcagctgg	cacgacaggt	ttccccgactg	gaaagcgggc	agtgagcgca	acgcaattaa	2520
tgtgagttag	ctcactcatt	aggcacccca	ggctttacac	tttatgcttc	cggtctgtat	2580
gttgtgtgga	attgtgagcg	gataacaatt	tcacacagga	aacagctatg	accatgatta	2640
cgccaagctt	gcatgcctgc	aggtcgactc	tagaggatcc	ccgcattgcg	gattacgtat	2700
tctaattgtc	agataacttc	gtatagcata	cattatacga	agttatctag	ggattcataa	2760
ccattttctc	aatcgaatta	cacagaacac	accgtacaaa	cctctctatc	ataactactt	2820
aatagtcaca	cacgtactcg	tctaaataca	catcatcgtc	ctacaagttc	atcaaagtgt	2880
tggacagaca	actataccag	catggatctc	ttgtatcggt	tcttttctcc	cgctctctcg	2940
caataacaat	gaacactggg	tcaatcatag	cctacacagg	tgaacagagt	agcgtttata	3000
cagggtttat	acggtgattc	ctacggcaaa	aatttttcat	ttctaataaaa	aaaaagaaaa	3060
atttttcttt	ccaacgctag	aaggaaaaga	aaaatctaata	taaattgatt	tggtgatttt	3120
ctgagagttc	cctttttcat	atatcgaatt	ttgaatataa	aaggagatcg	aaaaaatttt	3180
tctattcaat	ctgttttctg	gttttatttg	atagtttttt	tgtgtattat	tattatggat	3240
tagtactggg	ttatatgggt	ttttctgtat	aacttctttt	tatttttagtt	tgtttaatct	3300
tattttgagt	tacattatag	ttccctaact	gcaagagaag	taacattaaa	actcgagatg	3360
ggtaaggaaa	agactcacgt	ttcgaggccg	cgattaaatt	ccaacatgga	tgctgattta	3420
tatgggtata	aatgggctcg	cgataatgtc	gggcaatcag	gtgcgacaat	ctatcgattg	3480
tatgggaagc	ccgatgcgcc	agagttgttt	ctgaaacatg	gcaaaggtag	cgttgccaat	3540
gatgttacag	atgagatggg	cagactaaac	tggctgacgg	aatttatgcc	tcttccgacc	3600
atcaagcatt	ttatccgtac	tcctgatgat	gcatggttac	tcaccactgc	gatccccggc	3660
aaaacagcat	tccaggtatt	agaagaatat	cctgattcag	gtgaaaatat	tgttgatgcg	3720
ctggcagtgt	tcctgcgccg	gttgcatctg	attcctgttt	gtaattgtcc	ttttaacagc	3780
gatcgcgat	ttcgtctcgc	tcaggcgcaa	tcacgaatga	ataacggttt	ggttgatgcg	3840
agtgattttg	atgacgagcg	taatggctgg	cctgttgaac	aagtctggaa	agaaatgcat	3900
aagcttttgc	cattctcacc	ggattcagtc	gtcactcatg	gtgatttctc	acttgataac	3960
cttatttttg	acgaggggaa	attaataggt	tgtattgatg	ttggacgagt	cggaatcgca	4020
gaccgatacc	aggatcttgc	catcctatgg	aactgcctcg	gtgagttttc	tccttcatta	4080
cagaaacggc	tttttcaaaa	atatggtatt	gataatcctg	atatgaataa	attgcagttt	4140
catttgatgc	tcgatgagtt	tttctaagtt	taacttgata	ctactagatt	ttttctcttc	4200
atttataaaa	tttttggtta	taattgaagc	tttagaagta	tgaaaaaatc	cttttttttc	4260
attctttgca	accaaataa	gaagcttctt	ttattcattg	aaatgatgaa	tataaaccta	4320
acaaaagaaa	aagactcgaa	tatcaaacat	taaaaaaaaa	taaaagaggt	tatctgtttt	4380
cccatttagt	tggagtttgc	attttctaata	agatagaact	ctcaattaat	gtggatttag	4440
tttctctgtt	cgtttttttt	tgttttgttc	tcactgtatt	tacatttcta	tttagtattt	4500

agttattcat ataatctata acttcgtata gcatacatta tacgaagtta tccagtgatg 4560
 atacaacgag ttagccaagg tg 4582

<210> 130
 <211> 80
 <212> DNA
 <213> Arti f i c i a l Sequence

<220>
 <223> BK505 primer

<400> 130
 ttccggtttc tttgaaattt ttttgattcg gtaatctccg agcagaagga gcattgcgga 60
 ttacgtattc taatgttcag 80

<210> 131
 <211> 81
 <212> DNA
 <213> Arti f i c i a l Sequence

<220>
 <223> BK506 primer

<400> 131
 gggtaataac tgatataatt aaattgaagc tctaatttgt gagtttagta caccttggct 60
 aactcgttgt atcatcactg g 81

<210> 132
 <211> 38
 <212> DNA
 <213> Arti f i c i a l Sequence

<220>
 <223> LA468 primer

<400> 132
 gcctcgagtt ttaatgttac ttctcttgca gttaggga 38

<210> 133
 <211> 31
 <212> DNA
 <213> Arti f i c i a l Sequence

<220>
 <223> LA492 primer

<400> 133
 gctaaattcg agtgaaacac aggaagacca g 31

<210> 134
 <211> 23
 <212> DNA
 <213> Arti f i c i a l Sequence

<220>
 <223> AK109-1 primer

<400> 134
 agtcacatca agatcgttta tgg 23

<210> 135
 <211> 23
 <212> DNA
 <213> Arti f i c i a l S e q u e n c e

 <220>
 <223> AK109-2 primer

 <400> 135
 gcacggaata tgggactact tcg 23

<210> 136
 <211> 23
 <212> DNA
 <213> Arti f i c i a l S e q u e n c e

 <220>
 <223> AK109-3 primer

 <400> 136
 actccacttc aagtaagagt ttg 23

<210> 137
 <211> 24
 <212> DNA
 <213> Arti f i c i a l S e q u e n c e

 <220>
 <223> oBP452 primer

 <400> 137
 ttctcgacgt gggccttttt cttg 24

<210> 138
 <211> 49
 <212> DNA
 <213> Arti f i c i a l S e q u e n c e

 <220>
 <223> oBP453 primer

 <400> 138
 tgcagcttta aataatcggg gtcactactt tgccttcggt tatcttgcc 49

<210> 139
 <211> 49
 <212> DNA
 <213> Arti f i c i a l S e q u e n c e

 <220>
 <223> oBP454 primer

 <400> 139
 gagcaggcaa gataaacgaa ggcaaagtag tgacaccgat tattttaaag 49

<210> 140
 <211> 49
 <212> DNA
 <213> Arti f i c i a l S e q u e n c e

 <220>

<223> oBP455 primer
 <400> 140
 tatggaccct gaaaccacag ccacattgta accaccacga cggttgttg 49

 <210> 141
 <211> 49
 <212> DNA
 <213> Arti f i c i a l Sequence

 <220>
 <223> oBP456 primer
 <400> 141
 tttagcaaca accgtcgtgg tggttacaat gtggctgtgg tttcagggt 49

 <210> 142
 <211> 49
 <212> DNA
 <213> Arti f i c i a l Sequence

 <220>
 <223> oBP457 primer
 <400> 142
 ccagaaacc tatacctgtg tggacgtaag gccatgaagc tttttcttt 49

 <210> 143
 <211> 49
 <212> DNA
 <213> Arti f i c i a l Sequence

 <220>
 <223> oBP458 primer
 <400> 143
 attggaaaga aaaagcttca tggccttacg tccacacagg tatagggtt 49

 <210> 144
 <211> 22
 <212> DNA
 <213> Arti f i c i a l Sequence

 <220>
 <223> oBP459 primer
 <400> 144
 cataagaaca cctttggtgg ag 22

 <210> 145
 <211> 22
 <212> DNA
 <213> Arti f i c i a l Sequence

 <220>
 <223> oBP460 primer
 <400> 145
 aggattatca ttcataagtt tc 22

 <210> 146
 <211> 20

<212> DNA
 <213> Arti fi ci al Sequence
 <220>
 <223> LA135 primer
 <400> 146
 cttggcagca acaggactag 20

<210> 147
 <211> 23
 <212> DNA
 <213> Arti fi ci al Sequence
 <220>
 <223> oBP461 primer
 <400> 147
 ttcttgagc tgggacatgt ttg 23

<210> 148
 <211> 22
 <212> DNA
 <213> Arti fi ci al Sequence
 <220>
 <223> LA92 primer
 <400> 148
 gagaagatgc ggccagcaaa ac 22

<210> 149
 <211> 4242
 <212> DNA
 <213> Arti fi ci al Sequence
 <220>
 <223> pLA59 pl asmi d
 <400> 149
 aaacgccagc aacgcggcct ttttacgggt cctggccttt tgctggcctt ttgctcacat 60
 gttctttcct gcgttatccc ctgattctgt ggataaccgt attaccgcct ttgagtgagc 120
 tgataccgct cgccgcagcc gaacgaccga gcgcagcgag tcagtgagcg aggaagcgga 180
 agagcgccca atacgcaaac cgcctctccc cgcgcggttg ccgattcatt aatgcagctg 240
 gcacgacagg tttcccgact ggaaagcggg cagtgagcgc aacgcaatta atgtgagtta 300
 gctcactcat taggcacccc aggctttaca ctttatgctt ccggctcgta tgttgtgtgg 360
 aattgtgagc ggataacaat ttcacacagg aaacagctat gaccatgatt acgccaagct 420
 tgcattgcctg caggtcgact ctagaggatc cgcaatgcgg atccgcattg cggattacgt 480
 attctaattg tcagtaccgt tcgtataatg tatgctatac gaagttatgc agattgtact 540
 gagagtgcac cataccacct tttcaattca tcattttttt ttatttcttt tttttgattt 600
 cggtttcctt gaaatttttt tgattcggta atctccgaac agaaggaaga acgaaggaag 660
 gagcacagac ttagattggg atatatacgc atatgtagtg ttgaagaaac atgaaattgc 720
 ccagtattct taacccaact gcacagaaca aaaacctgca ggaaacgaag ataaatcatg 780

20130927_CL5646WOPCT_ST25. txt

tcgaaagcta	catataagga	acgtgctgct	actcatccta	gtcctgttgc	tgccaagcta	840
tttaatatca	tgcacgaaaa	gcaaacaaac	ttgtgtgctt	cattggatgt	tcgtaccacc	900
aaggaattac	tggagttagt	tgaagcatta	ggccccaaaa	tttgtttact	aaaaacacat	960
gtggatatct	tgactgattt	ttccatggag	ggcacagtta	agccgctaaa	ggcattatcc	1020
gccaagtaca	atTTTTtact	cttcgaagac	agaaaatttg	ctgacattgg	taatacagtc	1080
aaattgcagt	actctgcggg	tgtatacaga	atagcagaat	gggcagacat	tacgaatgca	1140
cacggtgtgg	tgggccagg	tattgttagc	ggtttgaagc	aggcggcaga	agaagtaaca	1200
aaggaacct	gaggcctttt	gatgttagca	gaattgtcat	gcaagggctc	cctatctact	1260
ggagaatata	ctaagggtac	tgttgacatt	gcgaagagcg	acaaagattt	tgttatcggc	1320
tttattgctc	aaagagacat	gggtggaaga	gatgaaggtt	acgattggtt	gattatgaca	1380
cccgtgtgg	gttttagatga	caaggagac	gcattgggtc	aacagtatag	aaccgtggat	1440
gatgtggtct	ctacaggatc	tgacattatt	attgttggaa	gaggactatt	tgcaaagggga	1500
agggatgcta	aggtagaggg	tgaacgttac	agaaaagcag	gctgggaagc	atatttgaga	1560
agatgcggcc	agcaaaacta	aaaaactgta	ttataagtaa	atgcatgtat	actaaactca	1620
caaattagag	cttcaattta	attatatcag	ttattaccct	atgcggtgtg	aaataccgca	1680
cagatgcgta	aggagaaaat	accgcatcag	gaaattgtaa	acgttaatat	tttgttaaaa	1740
ttcgcgttaa	atTTTTgtta	aatcagctca	ttttttaacc	aataggccga	aatcggaaca	1800
atcccttata	aatcaaaaga	atagaccgag	atagggttga	gtgttggtcc	agtttggaac	1860
aagagtccac	tattaaagaa	cgtggactcc	aacgtcaaag	ggcgaaaaac	cgtctatcag	1920
ggcgatggcc	cactacgtga	accatcaccc	taatcaagat	aacttcgtat	aatgtatgct	1980
atacgaacgg	taccagtgat	gatacaacga	gttagccaag	gtgaattcac	tggccgtcgt	2040
tttacaacgt	cgtgactggg	aaaaccctgg	cgttacccaa	cttaatcgcc	ttgcagcaca	2100
tccccctttc	gccagctggc	gtaatagcga	agaggcccg	accgatcgcc	cttcccaaca	2160
gttgcgcagc	ctgaatggcg	aatggcgctt	gatgcggtat	tttctcctta	cgcatctgtg	2220
cggtatttca	caccgcatat	ggtgcactct	cagtacaatc	tgctctgatg	ccgcatagtt	2280
aagccagccc	cgacacccgc	caacacccgc	tgacgcgccc	tgacgggctt	gtctgctccc	2340
ggcatccgct	tacagacaag	ctgtgaccgt	ctccgggagc	tgcatgtgtc	agaggttttc	2400
accgtcatca	ccgaaacg	cgagacgaaa	gggcctcgtg	atacgcctat	ttttataggt	2460
taatgtcatg	ataataatgg	tttcttagac	gtcaggtggc	acttttcggg	gaaatgtgcg	2520
cggaaaccct	atttgtttat	ttttctaaat	acattcaa	atgtatccgc	tcatgagaca	2580
ataaccctga	taaatgcttc	aataatattg	aaaaaggaag	agtatgagta	ttcaacattt	2640
ccgtgtcgcc	cttattccct	tttttgccgc	atTTTgcctt	cctgtttttg	ctcaccacga	2700
aacgctgggtg	aaagtaaaag	atgctgaaga	tcagttgggt	gcacgagtg	gttacatcga	2760
actggatctc	aacagcggta	agatccttga	gagttttcgc	cccgaagaac	gttttccaat	2820

20130927_CL5646WOPCT_ST25. txt

gatgagcact tttaaagtgc tgctatgtgg cgcggtatta tcccgtattg acgccgggca 2880
agagcaactc ggtcgccgca tacactattc tcagaatgac ttgggttgagt actcaccagt 2940
cacagaaaag catcttacgg atggcatgac agtaagagaa ttatgcagtg ctgccataac 3000
catgagtgat aacactgcgg ccaacttact tctgacaacg atcggaggac cgaaggagct 3060
aaccgctttt ttgcacaaca tgggggatca tgtaactcgc cttgatcgtt gggaaccgga 3120
gctgaatgaa gccataccaa acgacgagcg tgacaccacg atgcctgtag caatggcaac 3180
aacgttgcg aaactattaa ctggcgaact acttactcta gcttcccggc aacaattaat 3240
agactggatg gaggcggata aagttgcagg accacttctg cgctcggccc ttccggctgg 3300
ctggtttatt gctgataaat ctggagccgg tgagcgtggg tctcgcggta tcattgcagc 3360
actggggcca gatggttaagc cctcccgtat cgtagttatc tacacgacgg ggagtcaggc 3420
aactatggat gaacgaaata gacagatcgc tgagataggt gcctcactga ttaagcattg 3480
gtaactgtca gaccaagttt actcatatat actttagatt gatttaaaac ttcattttta 3540
atttaaaagg atctaggtga agatcctttt tgataatctc atgacaaaaa tcccttaacg 3600
tgagttttcg ttccactgag cgtcagaccc cgtagaaaag atcaaaggat cttcttgaga 3660
tccttttttt ctgcgcgtaa tctgctgctt gcaaacaaaa aaaccaccgc taccagcggg 3720
ggtttgtttg ccgatcaag agctaccaac tctttttccg aaggtaactg gcttcagcag 3780
agcgagata ccaaatactg tccttctagt gtagccgtag ttaggccacc acttcaagaa 3840
ctctgtagca ccgcctacat acctcgctct gctaatacctg ttaccagtgg ctgctgccag 3900
tggcgataag tcgtgtctta ccgggttgga ctcaagacga tagttaccgg ataaggcgca 3960
gcggtcgggc tgaacggggg gttcgtgcac acagcccagc ttggagcgaa cgacctacac 4020
cgaactgaga tacctacagc gtgagctatg agaaagcgcc acgcttcccg aaggagagaaa 4080
ggcggacagg tatccggtaa gcggcagggt cggaacagga gagcgcacga gggagcttcc 4140
aggggggaaac gcctgggtatc tttatagtcc tgtcgggttt cgccacctct gacttgagcg 4200
tcgatttttg tgatgctcgt caggggggcg gagcctatgg aa 4242

<210> 150
<211> 80
<212> DNA
<213> Arti f i c i a l Sequence

<220>
<223> LA678 primer

<400> 150
caacgttaac accgttttcg gtttgccagg tgacttcaac ttgtccttgt gcattgcgga 60
ttacgtattc taatgttcag 80

<210> 151
<211> 81
<212> DNA
<213> Arti f i c i a l Sequence

<220>
 <223> LA679 primer

 <400> 151
 gtggagcatc gaagactggc aacatgattt caatcattct gatcttagag caccttggct 60
 aactcgttgt atcatcactg g 81

 <210> 152
 <211> 23
 <212> DNA
 <213> Arti fi ci al Sequence

 <220>
 <223> LA337 primer

 <400> 152
 ctcatttgaa tcagcttatg gtg 23

 <210> 153
 <211> 24
 <212> DNA
 <213> Arti fi ci al Sequence

 <220>
 <223> LA692 primer

 <400> 153
 ggaagtcatt gacaccatct tggc 24

 <210> 154
 <211> 24
 <212> DNA
 <213> Arti fi ci al Sequence

 <220>
 <223> LA693 primer

 <400> 154
 agaagctggg acagcagcgt tagc 24

 <210> 155
 <211> 7523
 <212> DNA
 <213> Arti fi ci al Sequence

 <220>
 <223> pLA34 pl asmi d

 <400> 155
 ccagcttttg ttcccttttag tgagggttaa ttgcgcgctt ggcgtaatca tggatcatagc 60
 tgtttcctgt gtgaaattgt tatccgctca caattccaca caacatagga gccggaagca 120
 taaagtgtaa agcctggggg gcctaataag tgaggtaact cacattaatt gcgttgcgct 180
 cactgcccgc tttccagtcg ggaaacctgt cgtgccagct gcattaatga atcggcccaac 240
 gcgcgggggag aggcggtttg cgtattgggc gctcttccgc ttcctcgctc actgactcgc 300
 tgcgctcggg cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg gtaatacggg 360
 tatccacaga atcaggggat aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg 420

20130927_CL5646WOPCT_ST25. txt

ccaggaaccg	taaaaaggcc	gcgttgctgg	cgtttttcca	taggctccgc	ccccctgacg	480
agcatcacia	aatcgacgc	tcaagtcaga	ggtagcgaaa	cccgacagga	ctataaagat	540
accaggcggt	tccccctgga	agctccctcg	tgcgctctcc	tgttccgacc	ctgccgctta	600
ccggatacct	gtccgccttt	ctcccttcgg	gaagcgtggc	gctttctcat	agctcacgct	660
gtaggtatct	cagttcgggt	taggtcggtc	gctccaagct	gggctgtgtg	cacgaacccc	720
ccgttcagcc	cgaccgctgc	gccttatccg	gtaactatcg	tcttgagtcc	aacccggtaa	780
gacacgactt	atcgccactg	gcagcagcca	ctggtaacag	gattagcaga	gcgaggatg	840
taggcggtgc	tacagagttc	ttgaagtggg	ggcctaacta	cggctacact	agaaggacag	900
tatttggtat	ctgcgctctg	ctgaagccag	ttaccttcgg	aaaaagagtt	ggtagctctt	960
gatccggcaa	acaaccacc	gctggtagcg	gtggtttttt	tgtttgcaag	cagcagatta	1020
cgcgagaaa	aaaaggatct	caagaagatc	ctttgatctt	ttctacgggg	tctgacgctc	1080
agtggaacga	aaactcacgt	taagggattt	tggtcatgag	attatcaaaa	aggatcttca	1140
cctagatcct	tttaaattaa	aatgaagtt	ttaaataaat	ctaaagtata	tatgagtaaa	1200
cttggtctga	cagttaccaa	tgcttaatca	gtgaggcacc	tatctcagcg	atctgtctat	1260
ttcgttcatc	catagttgcc	tgactccccg	tcgtgtagat	aactacgata	cgaggagggt	1320
taccatctgg	ccccagtgt	gcaatgatac	cgcgagaccc	acgctcaccg	gctccagatt	1380
tatcagcaat	aaaccagcca	gccggaaggg	ccgagcgag	aagtggctct	gcaactttat	1440
ccgcctccat	ccagtctatt	aattgttgcc	gggaagctag	agtaagtagt	tcgccagtta	1500
atagtttgcg	caacgttggt	gccattgcta	caggcatcgt	gggtgcacgc	tcgtcgtttg	1560
gtatggcttc	attcagctcc	ggttcccaac	gatcaaggcg	agttacatga	tccccatgt	1620
tgtgcaaaaa	agcggtagc	tccttcgggc	ctccgatcgt	tgtcagaagt	aagttggccg	1680
cagtgttatc	actcatgggt	atggcagcac	tgcataattc	tcttactgtc	atgccatccg	1740
taagatgctt	ttctgtgact	ggtgagtact	caaccaagtc	attctgagaa	tagtgtatgc	1800
ggcgaccgag	ttgctcttgc	ccggcgtaaa	tacgggataa	taccgagcca	catagcagaa	1860
ctttaaaagt	gctcatcatt	ggaaaacgtt	cttcggggcg	aaaactctca	aggatcttac	1920
cgctgttgag	atccagttcg	atgtaacca	ctcgtgcacc	caactgatct	tcagcatctt	1980
ttactttcac	cagcgtttct	gggtgagcaa	aaacaggaag	gcaaaatgcc	gcaaaaaagg	2040
gaataagggc	gacacggaaa	tgttgaaata	tcatactctt	cctttttcaa	tattattgaa	2100
gcatttatca	gggttattgt	ctcatgagcg	gatacatatt	tgaatgtatt	tagaaaaata	2160
aacaaatagg	ggttccgcgc	acatttcccc	gaaaagtgcc	acctgaacga	agcatctgtg	2220
cttcattttg	tagaacaata	atgcaacgcg	agagcgctaa	tttttcaaac	aaagaatctg	2280
agctgcattt	ttacagaaca	gaaatgcaac	gcgaaagcgc	tattttacca	acgaagaatc	2340
tgtgcttcat	ttttgtaaaa	caaaaatgca	acgcgagagc	gctaattttt	caaacaaga	2400
atctgagctg	catttttaca	gaacagaaat	gcaacgcgag	agcgctatct	taccaacaaa	2460

20130927_CL5646WOPCT_ST25. txt

gaatctatac	ttcttttttg	ttctacaaaa	atgcatcccg	agagcgctat	ttttctaaca	2520
aagcatctta	gattactttt	tttctccttt	gtgcgctcta	taatgcagtc	tcttgataac	2580
tttttgcact	gtaggtccgt	taaggttaga	agaaggctac	tttggtgtct	attttctctt	2640
ccataaaaaa	agcctgactc	cacttcccgc	gtttactgat	tactagcgaa	gctgcgggtg	2700
cattttttca	agataaaggc	atccccgatt	atattctata	ccgatgtgga	ttgcgcatac	2760
tttgtgaaca	gaaagtgata	gcgttgatga	ttcttcattg	gtcagaaaaat	tatgaacggt	2820
ttcttctatt	ttgtctctat	atactacgta	taggaaatgt	ttacattttc	gtattgtttt	2880
cgattcactc	tatgaatagt	tcttactaca	atttttttgt	ctaaagagta	atactagaga	2940
taaacataaa	aaatgtagag	gtcgagttaa	gatgcaagtt	caaggagcga	aagggtggatg	3000
ggtaggttat	atagggatat	agcacagaga	tatatagcaa	agagatactt	ttgagcaatg	3060
tttgtggaag	cggtattcgc	aatattttag	tagctcgtaa	cagtcgggtg	cgtttttgggt	3120
tttttgaaag	tgcgtcttca	gagcgctttt	ggttttcaaa	agcgctctga	agttcctata	3180
ctttctagag	aataggaact	tcggaatagg	aacttcaaag	cgtttccgaa	aacgagcgct	3240
tccgaaaatg	caacgcgagc	tgcgcacata	cagctcactg	ttcacgtcgc	acctatatct	3300
gcgtgttgcc	tgtatatata	tatacatgag	aagaacggca	tagtgcgtgt	ttatgcttaa	3360
atgcgtactt	atatgcgtct	atttatgtag	gatgaaagggt	agtctagtag	ctcctgtgat	3420
attatcccat	tccatgcggg	gtatcgtatg	cttccttcag	cactaccctt	tagctgttct	3480
atatgctgcc	actcctcaat	tggattagtc	tcattccttca	atgctatcat	ttcctttgat	3540
attggatcat	ctaagaaacc	attattatca	tgacattaac	ctataaaaaat	aggcgtatca	3600
cgaggccctt	tcgtctcgcg	cgtttcgggtg	atgacggtga	aaacctctga	cacatgcagc	3660
tcccgagagac	ggtcacagct	tgtctgtaag	cggatgccgg	gagcagacaa	gcccgtcagg	3720
gcgcgtcagc	gggtgttggc	gggtgtcggg	gctggcttaa	ctatgcggca	tcagagcaga	3780
ttgtactgag	agtgacccat	aaattcccgt	tttaagagct	tgggtgagcgc	taggagtcac	3840
tgccaggat	cgtttgaaca	cggcattagt	caggggaagtc	ataacacagt	cctttcccgc	3900
aattttcttt	ttctattact	cttggcctcc	tctagtagac	tctatatattt	tttatgcctc	3960
ggtaatgatt	ttcatttttt	tttttcccct	agcggtatgac	tctttttttt	tcttagcgat	4020
tggcattatc	acataatgaa	ttatacatta	tataaagtaa	tgtgatttct	tcgaagaata	4080
tactaaaaaa	tgagcaggca	agataaacga	aggcaaagat	gacagagcag	aaagccctag	4140
taaagcgtat	tacaaatgaa	accaagattc	agattgcgat	ctcttttaaag	ggtggtcccc	4200
tagcgataga	gcactcgatc	ttcccagaaa	aagaggcaga	agcagtagca	gaacaggcca	4260
cacaatcgca	agtgattaac	gtccacacag	gtataggggtt	tctggaccat	atgatacatg	4320
ctctggccaa	gcattccggc	tggctcgctaa	tcgttgagtg	cattggtgac	ttacacatag	4380
acgaccatca	caccactgaa	gactgcggga	ttgctctcgg	tcaagctttt	aaagaggccc	4440
tactggcgcg	tggagtaaaa	aggtttggat	caggattttgc	gcctttggat	gaggcacttt	4500

20130927_CL5646WOPCT_ST25. txt

ccagagcggg	ggtagatctt	tcgaacaggc	cgtacgcagt	tgtcgaactt	ggtttgcaaa	4560
gggagaaagt	aggagatctc	tcttgcgaga	tgatcccgc	ttttcttgaa	agctttgcag	4620
aggctagcag	aattaccctc	cacgttgatt	gtctgcgagg	caagaatgat	catcaccgta	4680
gtgagagtgc	gttcaaggct	cttgcggttg	ccataagaga	agccacctcg	cccaatggta	4740
ccaacgatgt	tccctccacc	aaagggtgtt	ttatgtagt	acaccgatta	tttaaagctg	4800
cagcatatga	tatatataca	tgtgtatata	tgtataccta	tgaatgtcag	taagtatgta	4860
tacgaacagt	atgatactga	agatgacaag	gtaatgcata	attctatacg	tgtcattctg	4920
aacgaggcgc	gctttccttt	tttctttttg	ctttttcttt	ttttttctct	tgaactcgac	4980
ggatctatgc	ggtgtgaaat	accgcacaga	tgcgtaagga	gaaaataaccg	catcaggaaa	5040
ttgtaaactg	taatatattg	ttaaaattcg	cgtaaatttt	ttgttaaata	agctcatttt	5100
ttaaccaata	ggccgaaatc	ggcaaaatcc	cttataaatc	aaaagaatag	accgagatag	5160
ggttgagtgt	tgttccagtt	tggaacaaga	gtccactatt	aaagaacgtg	gactccaacg	5220
tcaaagggcg	aaaaaccgtc	tatcagggcg	atggcccact	acgtgaacca	tcaccctaata	5280
caagtttttt	ggggctcgagg	tgccgtaaa	cactaaatcg	gaaccctaaa	gggagccccc	5340
gatttagagc	ttgacgggga	aagccggcga	acgtggcgag	aaaggaaggg	aagaaagcga	5400
aaggagcggg	cgctagggcg	ctggcaagt	tagcgggtac	gctgcgcgta	accaccacac	5460
ccgccgcgct	taatgcgccg	ctacagggcg	cgtcgcgcc	ttcgccattc	aggctgcgca	5520
actgttggga	agggcgatcg	gtgcgggcct	cttcgctatt	acgccagctg	gcgaaagggg	5580
gatgtgctgc	aaggcgatta	agttgggtaa	cgccagggtt	ttcccagtca	cgacgttgta	5640
aaacgacggc	cagtgcgcgc	gcgtaatacg	actcactata	gggcgaattg	ggtaccgggc	5700
ccccctcga	ggtattagaa	gccgccgagc	gggcgacagc	cctccgacgg	aagactctcc	5760
tccgtgcgtc	ctcgtcttca	ccggtcgcgt	tcctgaaacg	cagatgtgcc	tcgcgccgca	5820
ctgctccgaa	caataaagat	tctacaatac	tagcttttat	ggttatgaag	aggaaaaatt	5880
ggcagtaacc	tggccccaca	aaccttcaaa	ttaacgaatc	aaattaacaa	ccataggatg	5940
ataatgcgat	tagtttttta	gccttatttc	tggggtaatt	aatcagcgaa	gcgatgattt	6000
ttgatctatt	aacagatata	taaatgaaa	agctgcataa	ccactttaac	taatactttc	6060
aacattttca	gtttgtatta	cttcttattc	aaatgtcata	aaagtatcaa	caaaaaattg	6120
ttaatatacc	tctatacttt	aacgtcaagg	agaaaaatgt	ccaatttact	gcccgtacac	6180
caaaatttgc	ctgcattacc	ggtcgatgca	acgagtgatg	aggttcgcaa	gaacctgatg	6240
gacatgttca	gggatcgcca	ggcgttttct	gagcatacct	ggaaaatgct	tctgtccggt	6300
tgccggctcg	gggcggcatg	gtgcaagttg	aataaccgga	aatggttttc	cgcagaacct	6360
gaagatgttc	gcgattatct	tctatatctt	caggcgcgcg	gtctggcagt	aaaaactatc	6420
cagcaacatt	tgggccagct	aaacatgctt	catcgtcggt	ccgggctgcc	acgaccaagt	6480
gacagcaatg	ctgtttcact	ggttatgcgg	cggatccgaa	aagaaaacgt	tgatgccggt	6540

20130927_CL5646WOPCT_ST25. txt

gaacgtgcaa aacaggctct agcgttcgaa cgcactgatt tcgaccaggt tcgttcactc	6600
atggaaaata gcgatcgctg ccaggatata cgtaatctgg cttttctggg gattgcttat	6660
aacaccctgt tacgtatagc cgaaattgcc aggatcaggg ttaaagatat ctcacgtact	6720
gacggtggga gaatgttaat ccatattggc agaacgaaaa cgctggttag caccgcaggt	6780
gtagagaagg cacttagcct gggggtaact aaactggtcg agcgatggat ttccgtctct	6840
ggtgtagctg atgatccgaa taactacctg ttttgccggg tcagaaaaaa tgggtgttgc	6900
gcgccatctg ccaccagcca gctatcaact cgcgccctgg aagggtttt tgaagcaact	6960
catcgattga tttacggcgc taaggatgac tctggtcaga gatactggc ctggtctgga	7020
cacagtgcc gtgtcggagc cgcgcgagat atggcccgcg ctggagtttc aataccggag	7080
atcatgcaag ctggtggctg gaccaatgta aatattgtca tgaactatat ccgtaacctg	7140
gatagtgaag caggggcaat ggtgcgcctg ctggaagatg gcgattagga gtaagcgaat	7200
ttcttatgat ttatgatttt tattattaaa taagttataa aaaaaataag tgtatacaaa	7260
ttttaagtg actcttaggt tttaaaacga aaattcttat tcttgagtaa ctctttcctg	7320
taggtcaggt tgctttctca ggtatagcat gaggtcgctc ttattgacca cacctctacc	7380
ggcatgccga gcaaatgcct gcaaatcgct cccattttca cccaattgta gatatgctaa	7440
ctccagcaat gagttgatga atctcgggtg gtattttatg tcctcagagg acaacacctg	7500
tggtccgcca ccgcggtgga gct	7523

<210> 156
 <211> 96
 <212> DNA
 <213> Arti f i c i a l Sequence

<220>
 <223> LA722 pl asmi d

<400> 156	
tgccaattat ttacctaaac atctataacc ttcaaaagta aaaaaataca caaacgttga	60
atcatcacct tggctaactc gttgtatcat cactgg	96

<210> 157
 <211> 80
 <212> DNA
 <213> Arti f i c i a l Sequence

<220>
 <223> LA733 primer

<400> 157	
cataatcaat ctcaaagaga acaacacaat acaataacaa gaagaacaaa gcattgcgga	60
ttacgtattc taatgttcag	80

<210> 158
 <211> 30
 <212> DNA
 <213> Arti f i c i a l Sequence

<220>
 <223> LA453 primer
 <400> 158
 caccgaagaa gaatgcaaaa atttcagctc 30

<210> 159
 <211> 25
 <212> DNA
 <213> Arti f i c i a l Sequence
 <220>
 <223> LA694 primer
 <400> 159
 gctgaagttg ttagaactgt tgttg 25

<210> 160
 <211> 21
 <212> DNA
 <213> Arti f i c i a l Sequence
 <220>
 <223> LA695 primer
 <400> 160
 tgtagctgg agtagacttg g 21

<210> 161
 <211> 22
 <212> DNA
 <213> Arti f i c i a l Sequence
 <220>
 <223> oBP594 primer
 <400> 161
 agctgtctcg tggtgtgggt tt 22

<210> 162
 <211> 49
 <212> DNA
 <213> Arti f i c i a l Sequence
 <220>
 <223> oBP595 primer
 <400> 162
 cttaataata gaacaatc atcctttacg ggcatttat agtgtcgtt 49

<210> 163
 <211> 49
 <212> DNA
 <213> Arti f i c i a l Sequence
 <220>
 <223> oBP596 primer
 <400> 163
 gcgccaacga cactataaga tgcccgtaaa ggatgatatt gttctatta 49

<210> 164
 <211> 49
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> oBP597 primer

 <400> 164
 tatggaccct gaaaccacag ccacattgca acgacgacaa tgccaaacc 49

 <210> 165
 <211> 49
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> oBP598 primer

 <400> 165
 tccttggttt ggcattgtcg tcgttgcaat gtggctgtgg tttcagggt 49

 <210> 166
 <211> 49
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> oBP599 primer

 <400> 166
 atcctctcgc ggagtccttg ttcagtaaag gccatgaagc tttttcttt 49

 <210> 167
 <211> 49
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> oBP600 primer

 <400> 167
 attggaaga aaaagcttca tggcctttac tgaacaggga ctccgcgag 49

 <210> 168
 <211> 22
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> oBP601 primer

 <400> 168
 tcataccaca atcttagacc at 22

 <210> 169
 <211> 21
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> Primer oBP602

<400> 169
tggtcaaacc cctaaccaac c 21

<210> 170
<211> 22
<212> DNA
<213> Arti fi ci al Sequence

<220>
<223> oBP603 primer

<400> 170
tggtcccaca atctattacc ta 22

<210> 171
<211> 31
<212> DNA
<213> Arti fi ci al Sequence

<220>
<223> LA811 primer

<400> 171
aacgaagcat ctgtgcttca tttttagaa c 31

<210> 172
<211> 59
<212> DNA
<213> Arti fi ci al Sequence

<220>
<223> LA817 primer

<400> 172
cgatccactt gtatatattgg atgaattttt gaggaattct gaaccagtcc taaaacgag 59

<210> 173
<211> 31
<212> DNA
<213> Arti fi ci al Sequence

<220>
<223> LA812 primer

<400> 173
aacaaagata tgctattgaa gtgcaagatg g 31

<210> 174
<211> 33
<212> DNA
<213> Arti fi ci al Sequence

<220>
<223> LA818 primer

<400> 174
ctcaaaaatt catccaaata tacaagtgga tcg 33

<210> 175
<211> 90
<212> DNA
<213> Arti fi ci al Sequence

<220>
 <223> LA512 primer

 <400> 175
 gtattttggg agattcaatt ctctttccct ttccttttcc ttcgctcccc ttccttatca 60
 gcattgcgga ttacgtattc taatgttcag 90

 <210> 176
 <211> 90
 <212> DNA
 <213> Arti fi ci al Sequence

 <220>
 <223> LA513 primer

 <400> 176
 ttggttgggg gaaaaagagg caacaggaaa gatcagaggg ggaggggggg ggagagtgtc 60
 accttggtta actcgttgta tcatcactgg 90

 <210> 177
 <211> 29
 <212> DNA
 <213> Arti fi ci al Sequence

 <220>
 <223> LA516 primer

 <400> 177
 ctcgaaacaa taagacgacg atggctctg 29

 <210> 178
 <211> 30
 <212> DNA
 <213> Arti fi ci al Sequence

 <220>
 <223> LA514 primer

 <400> 178
 cactatctgg tgcaaacttg gcaccggaag 30

 <210> 179
 <211> 29
 <212> DNA
 <213> Arti fi ci al Sequence

 <220>
 <223> LA515 primer

 <400> 179
 tgttttagc cactcgtgaa cttctctgc 29

 <210> 180
 <211> 6903
 <212> DNA
 <213> Arti fi ci al Sequence

 <220>
 <223> pLA71 pl asmi d

20130927_CL5646WOPCT_ST25. txt

```

<400> 180
aacgccagc aacgcggcct ttttacggtt cctggccttt tgctggcctt ttgctcacat      60
gttctttcct gcgttatccc ctgattctgt ggataaccgt attaccgcct ttgagtgagc      120
tgataccgct cgccgcagcc gaacgaccga gcgcagcgag tcagtgagcg aggaagcgga      180
agagcgccca atacgcaaac cgcctctccc cgcgcgttgg ccgattcatt aatgcagctg      240
gcacgacagg tttcccgact ggaaagcggg cagtgagcgc aacgcaatta atgtgagtta      300
gctcactcat taggcacccc aggctttaca ctttatgctt ccggctcgta tgttgtgtgg      360
aattgtgagc ggataacaat ttcacacagg aaacagctat gaccatgatt acgccaagct      420
tgcattcgat ctgaaatgaa taacaatact gacagtagat ctgaaatgaa taacaatact      480
gacagtacta aataattgcc tacttggtt cacatacggt gcatacgtcg atatagataa      540
taatgataat gacagcagga ttatcgtaat acgtaatagt tgaaaatctc aaaaatgtgt      600
gggtcattac gtaaataatg ataggaatgg gattcttcta tttttccttt ttccattcta      660
gcagccgtcg ggaaaacgtg gcatcctctc tttcgggctc aattggagtc acgctgccgt      720
gagcatcctc tctttcata tctaacaact gagcacgtaa ccaatggaaa agcatgagct      780
tagcgttgct ccaaaaaagt attggatggg taataccatt tgtctgttct cttctgactt      840
tgactcctca aaaaaaaaaa atctacaatc aacagatcgc ttcaattacg ccctcacaaa      900
aacttttttc cttcttcttc gcccacgtta aattttatcc ctcatgttgt ctaacggatt      960
tctgcacttg atttattata aaaagacaaa gacataatac ttctctatca atttcagtta     1020
ttgttcttcc ttgcgttatt cttctgttct tctttttctt ttgtcatata taaccataac     1080
caagtaatac atattcaaat ctagagctga ggatgttgac aaaagcaaca aaagaacaaa     1140
aatcccttgt gaaaaacaga ggggcggagc ttgttgttga ttgcttagtg gagcaagggtg     1200
tcacacatgt atttggcatt ccaggtgcaa aaattgatgc ggtatttgac gctttacaag     1260
ataaaggacc tgaaattatc gttgcccggc acgaacaaaa cgcagcattc atggcccaag     1320
cagtcggccg tttaactgga aaaccgggag tcgtgttagt cacatcagga ccgggtgcct     1380
ctaacttggc aacaggcctg ctgacagcga aactgaagg agaccctgtc gttgcgcttg     1440
ctggaaacgt gatccgtgca gatcgtttaa aacggacaca tcaatctttg gataatgcgg     1500
cgctattcca gccgattaca aaatacagtg tagaagttca agatgtaaaa aatataccgg     1560
aagctgttac aaatgcattt aggatagcgt cagcagggca ggctggggcc gcttttgtga     1620
gctttccgca agatgttgtg aatgaagtca caaatacgaa aaacgtgcgt gctgttgcag     1680
cgccaaaact cggctcctga gcagatgatg caatcagtgc ggccatagca aaaatccaaa     1740
cagcaaaaact tcctgtcgtt ttggtcggca tgaaaggcgg aagaccggaa gcaattaaag     1800
cggttcgcaa gcttttgaaa aagggttcagc ttccatttgt tgaaacatat caagctgccg     1860
gtaccctttc tagagattta gaggatcaat attttgcccg tatcggtttg ttccgcaacc     1920
agcctggcga tttactgcta gagcaggcag atgttgttct gacgatcggc tatgaccgga     1980
ttgaatatga tccgaaattc tggaatatca atggagaccg gacaattatc catttagacg     2040

```

20130927_CL5646WOPCT_ST25. txt

agattatcgc	tgacattgat	catgcttacc	agcctgatct	tgaattgatc	ggtgacattc	2100
cgtccacgat	caatcatatc	gaacacgatg	ctgtgaaagt	ggaatttgca	gagcgtgagc	2160
agaaaatcct	ttctgattta	aaacaatata	tgcatagaagg	tgagcaggtg	cctgcagatt	2220
ggaaatcaga	cagagcgcac	cctcttgaaa	tcgttaaaga	gttgcgtaat	gcagtcgatg	2280
atcatgttac	agtaacttgc	gatatcgggt	cgcacgccat	ttggatgtca	cgttattttcc	2340
gcagctacga	gccgttaaca	ttaatgatca	gtaacgggat	gcaaacactc	ggcgttgccg	2400
ttccttgggc	aatcggcgct	tcattgggtga	aaccgggaga	aaaagtgggt	tctgtctctg	2460
gtgacggcgg	tttcttattc	tcagcaatgg	aattagagac	agcagttcga	ctaaaagcac	2520
caattgtaca	cattgtatgg	aacgacagca	catatgacat	ggttgcattc	cagcaattga	2580
aaaaatataa	ccgtacatct	gcggtcgatt	tcggaaatat	cgatatcgtg	aaatatgcgg	2640
aaagcttcgg	agcaactggc	ttgcgcgtag	aatcaccaga	ccagctggca	gatgttctgc	2700
gtcaaggcat	gaacgctgaa	ggtcctgtca	tcatcgatgt	cccggttgac	tacagtgata	2760
acattaattt	agcaagtgc	aagcttccga	aagaattcgg	ggaactcatg	aaaacgaaag	2820
ctctctagtt	aattaatcat	gtaattagtt	atgtcacgct	tacattcacg	ccctcccccc	2880
acatccgctc	taaccgaaaa	ggaaggagtt	agacaacctg	aagtctaggt	ccctattttat	2940
ttttttatag	ttatgttagt	attaagaacg	ttattttatat	ttcaaatttt	tctttttttt	3000
ctgtacagac	gcgtgtacgc	atgtaacatt	atactgaaaa	ccttgcttga	gaaggttttg	3060
ggacgctcga	aggctttaat	ttaggttttg	ggacgctcga	aggctttaat	ttggatccgc	3120
attgcggatt	acgtattcta	atgttcagta	ccgttcgtat	aatgtatgct	atacgaagtt	3180
atgcagattg	tactgagagt	gcaccatacc	acagcttttc	aattcaattc	atcatttttt	3240
ttttattctt	ttttttgatt	tcggtttctt	tgaaattttt	ttgattcggt	aatctccgaa	3300
cagaaggaag	aacgaaggaa	ggagcacaga	cttagattgg	tatatatacg	catatgtagt	3360
gttgaagaaa	catgaaattg	cccagtattc	ttaacccaac	tgacacagaac	aaaaacctgc	3420
aggaaacgaa	gataaatcat	gtcgaaagct	acataataagg	aacgtgctgc	tactcatcct	3480
agtcctgttg	ctgccaagct	atttaatatc	atgcacgaaa	agcaaacaaa	cttgtgtgct	3540
tcattggatg	ttcgtaccac	caaggaatta	ctggagttag	ttgaagcatt	aggtcccaaa	3600
atttgtttac	taaaaacaca	tgtggatatc	ttgactgatt	tttccatgga	gggcacagtt	3660
aagccgctaa	aggcattatc	cgccaagtac	aatttttttac	tcttcgaaga	cagaaaattt	3720
gctgacattg	gtaatacagt	caaattgcag	tactctgcgg	gtgtatacag	aatagcagaa	3780
tgggcagaca	ttacgaatgc	acacggtgtg	gtgggccag	gtattgttag	cggtttgaag	3840
caggcggcag	aagaagtaac	aaaggaacct	agaggccctt	tgatgttagc	agaattgtca	3900
tgcaagggct	ccctatctac	tggagaatat	actaagggtg	ctgttgacat	tgcaagagc	3960
gacaaagatt	ttgttatcgg	ctttattgct	caaagagaca	tgggtggaag	agatgaaggt	4020
tacgattggg	tgattatgac	acccggtgtg	ggtttagatg	acaagggaga	cgcattgggt	4080

20130927_CL5646WOPCT_ST25. txt

caacagtata	gaaccgtgga	tgatgtggtc	tctacaggat	ctgacattat	tattgttggg	4140
agaggactat	ttgcaaaggg	aagggatgct	aaggtagagg	gtgaacgtta	cagaaaagca	4200
ggctgggaag	catatttgag	aagatgcggc	cagcaaaact	aaaaaactgt	attataagta	4260
aatgcatgta	tactaaactc	acaaattaga	gcttcaatth	aattatatca	gttattaccc	4320
tatgcggtgt	gaaataccgc	acagatgcgt	aaggagaaaa	taccgcatca	ggaaattgta	4380
aacgttaata	ttttgttaaa	attcgcgtta	aatthttgtt	aatcagctc	atthtttaac	4440
caataggccg	aaatcggcaa	aatcccttat	aatcaaaag	aatagaccga	gatagggttg	4500
agtgttgttc	cagtttgga	caagagtcca	ctattaaaga	acgtggactc	caacgtcaaa	4560
gggcgaaaaa	ccgtctatca	gggcgatggc	ccactacgtg	aaccatcacc	ctaataaga	4620
taacttcgta	taatgtatgc	tatacgaacg	gtaccagtga	tgatacaacg	agttagccaa	4680
ggtgaattca	ctggccgtcg	ttttacaacg	tcgtgactgg	gaaaaccctg	gcgttaccca	4740
acttaatcgc	cttgcagcac	atccccctth	cgccagctgg	cgtaatagcg	aagaggcccg	4800
caccgatcgc	ccttccaac	agttgcgcag	cctgaatggc	gaatggcgcc	tgatgcggta	4860
ttttctcctt	acgcatctgt	gcggtatthc	acaccgcata	tggtgcactc	tcagtacaat	4920
ctgctctgat	gccgcatagt	taagccagcc	ccgacacccg	ccaacacccg	ctgacgcgcc	4980
ctgacgggct	tgtctgctcc	cggcatccgc	ttacagacaa	gctgtgaccg	tctccgggag	5040
ctgcatgtgt	cagaggttht	caccgtcatc	accgaaacgc	gcgagacgaa	agggcctcgt	5100
gatacgccta	tttttatagg	ttaatgtcat	gataataatg	gtttcttaga	cgtcaggtgg	5160
cactthtcgg	ggaaatgtgc	gcggaacccc	tatttgttht	tttttctaaa	tacattcaaa	5220
tatgtatccg	ctcatgagac	aataaccctg	ataaatgctt	caataatatt	gaaaaaggaa	5280
gagtatgagt	attcaacatt	tccgtgtcgc	ccttattccc	ttttttgcgg	cattttgcct	5340
tcctgtthtt	gctcacccag	aaacgctgg	gaaagtaaaa	gatgctgaag	atcagttggg	5400
tgcacgagt	ggttacatcg	aactggatct	caacagcgg	aagatccttg	agagttthtcg	5460
ccccgaagaa	cgthttccaa	tgatgagcac	ttttaaagtt	ctgctatgtg	gcgcggtatt	5520
atcccgtatt	gacgccgggc	aagagcaact	cggctgcggc	atacactatt	ctcagaatga	5580
cttggttgag	tactcaccag	tcacagaaaa	gcattcttac	gatggcatga	cagtaagaga	5640
attatgcagt	gctgccataa	ccatgagtga	taacactgcg	gccaacttac	ttctgacaac	5700
gatcggagga	ccgaaggagc	taaccgctth	tttgacaaac	atgggggatc	atgtaactcg	5760
ccttgatcgt	tgggaaccgg	agctgaatga	agccatacca	aacgacgagc	gtgacaccac	5820
gatgcctgta	gcaatggcaa	caacgttgcg	caaactatta	actggcgaac	tacttactct	5880
agcttcccgg	caacaattaa	tagactggat	ggaggcggat	aaagttgcag	gaccacttct	5940
gcgctcggcc	cttccggctg	gctggthttat	tgctgataaa	tctggagccg	gtgagcgtgg	6000
gtctcgcggt	atcattgcag	cactggggcc	agatggtaag	ccctcccgt	tcgtagttat	6060
ctacacgacg	gggagtcagg	caactatgga	tgaacgaaat	agacagatcg	ctgagatagg	6120

tgctcactg attaagcatt ggtaactgtc agaccaagtt tactcatata tacttttagat	6180
tgatttaaaa cttcatTTTT aatttaaaag gatctaggtg aagatccttt ttgataatct	6240
catgaccaa atcccttaac gtgagttttc gttccactga gcgtcagacc ccgtagaaaa	6300
gatcaaagga tcttcttgag atcctTTTT tctgcgcgta atctgctgct tgcaaacaaa	6360
aaaaccaccg ctaccagcgg tggtttggtt gccggatcaa gagctaccaa ctctttttcc	6420
gaaggtaact ggcttcagca gagcgcagat accaaatact gtccttctag tgtagccgta	6480
gttaggccac cacttcaaga actctgtagc accgcctaca tacctcgctc tgctaatacct	6540
gttaccagtg gctgctgcc a gtggcgataa gtcgtgtctt accgggttg actcaagacg	6600
atagttaccg gataaggcgc agcggtcggg ctgaacgggg ggttcgtgca cacagcccag	6660
cttgagcga acgacctaca ccgaactgag atacctacag cgtgagctat gagaaagcgc	6720
cacgcttccc gaaggagaa aggcggacag gtatccggt a gcggcaggg tcggaacagg	6780
agagcgcacg agggagcttc cagggggaaa cgcctggtat ctttatagtc ctgtcgggtt	6840
tcgccacctc tgacttgagc gtcgattttt gtgatgctcg tcaggggggc ggagcctatg	6900
gaa	6903

<210> 181
 <211> 96
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> LA829 primer

<400> 181	
ccaaatttac aatatctcct gaattcttgg cttggaatat gggcagtaca gcttgtgtga	60
tattgcacct tggctaactc gttgtatcat cactgg	96

<210> 182
 <211> 90
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> LA834 primer

<400> 182	
atgtccaag gtagaaaagc tgcagaaaga ttggctaaga agactgtcct cattacaggt	60
gatctgaaat gaataacaat actgacagta	90

<210> 183
 <211> 29
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> N1257 primer

<400> 183	
gatgatgcta tttggtgcag aggggtgatg	29

<210> 184
 <211> 22
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> LA740 primer

<400> 184
 cgataatcct gctgtcatta tc 22

<210> 185
 <211> 29
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> LA830 primer

<400> 185
 cacggcaaac ttagaggcac aatagatag 29

<210> 186
 <211> 6924
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> pLA78 plasmid

<400> 186
 gatccgcatt gcggattacg tattctaattg ttcagtaccg ttcgtataat gtatgctata 60
 cgaagttatg cagattgtac tgagagtgc ccataccacc ttttcaattc atcatttttt 120
 ttttattctt ttttttgatt tcggtttcct tgaaattttt ttgattcggg aatctccgaa 180
 cagaaggaag aacgaaggaa ggagcacaga cttagattgg tatatatacg catatgtagt 240
 gttgaagaaa catgaaattg cccagtattc ttaacccaac tgcacagaac aaaaacctgc 300
 aggaaacgaa gataaatcat gtcgaaagct acatataagg aacgtgctgc tactcatcct 360
 agtcctgttg ctgccaagct atttaatatc atgcacgaaa agcaaacaaa cttgtgtgct 420
 tcattggatg ttcgtaccac caaggaatta ctggagttag ttgaagcatt aggtcccaaa 480
 atttgtttac taaaaacaca tgttgatatc ttgactgatt tttccatgga gggcacagtt 540
 aagccgctaa aggcattatc cgccaagtac aattttttac tcttcgaaga cagaaaattt 600
 gctgacattg gtaatacagt caaattgcag tactctgcgg gtgtatacag aatagcagaa 660
 tgggcagaca ttacgaatgc acacggtgtg gtgggccag gtattgttag cggtttgaag 720
 caggcggcag aagaagtaac aaaggaacct agaggccttt tgatgttagc agaattgtca 780
 tgcaagggct ccctatctac tggagaatat actaagggtg ctgttgacat tgcgaagagc 840
 gacaaagatt ttgttatcgg ctttattgct caaagagaca tgggtggaag agatgaaggt 900
 tacgattggg tgattatgac acccggtgtg ggttttagatg acaagggaga cgcattgggt 960
 caacagtata gaaccgtgga tgatgtgggtc tctacaggat ctgacattat tattgttgga 1020

20130927_CL5646WOPCT_ST25. txt

agaggactat	ttgcaaaggg	aagggatgct	aaggttagagg	gtgaacgtta	cagaaaagca	1080
ggctgggaag	catatttgag	aagatgcggc	cagcaaaact	aaaaaactgt	attataagta	1140
aatgcatgta	tactaaactc	acaaattaga	gcttcaatth	aattatatca	gttattaccc	1200
tatgcggtgt	gaaataccgc	acagatgcgt	aaggagaaaa	taccgcatca	ggaaattgta	1260
aacgttaata	ttttgttaaa	attcgcgtta	aatthttgtt	aatcagctc	atthtttaac	1320
caataggccg	aaatcggcaa	aatcccttat	aatcaaaaag	aatagaccga	gatagggttg	1380
agtgttgth	cagthttgaa	caagagtcca	ctattaaaga	acgtggactc	caacgtcaaa	1440
gggcgaaaaa	ccgtctatca	gggcgatggc	ccactacgtg	aaccatcacc	ctaatcaaga	1500
taacttcgta	taatgtatgc	tatacgaacg	gtaccagtga	tgatacaacg	agttagccaa	1560
ggtgaattca	ctggccgtcg	ttttacaacg	tcgtgactgg	gaaaaccctg	gcgttaccca	1620
acttaatcgc	cttgacgcac	atccccctth	cgccagctgg	cgtaatagcg	aagaggcccg	1680
caccgatcgc	ccttcccaac	agttgcgcag	cctgaatggc	gaatggcgcc	tgatgcggta	1740
ttttctcctt	acgcatctgt	gcggtatth	acaccgcata	tggtgcactc	tcagtacaat	1800
ctgctctgat	gccgcatagt	taagccagcc	ccgacaccgc	ccaacaccgc	ctgacgcgcc	1860
ctgacgggct	tgtctgctcc	cggcatccgc	ttacagacaa	gctgtgaccg	tctccgggag	1920
ctgcatgtgt	cagaggthtt	caccgtcatc	accgaaacgc	gcgagacgaa	agggcctcgt	1980
gatacgccta	tttttatagg	ttaatgtcat	gataataatg	gtttcttaga	cgtcagggtg	2040
cactthtcgg	ggaaatgtgc	gcggaacccc	tatthgtth	ttttctaaa	tacattcaaa	2100
tatgtatccg	ctcatgagac	aataaccctg	ataaatgctt	caataatatt	gaaaaaggaa	2160
gagtatgagt	attcaacatt	tccgtgtcgc	ccttattccc	ttttttgcgg	cattthgcct	2220
tcctgtthtt	gctcacccag	aaacgctgg	gaaagtaaaa	gatgctgaag	atcagttggg	2280
tgcacgagtg	ggttacatcg	aactggatct	caacagcgg	aagatccttg	agagthttcg	2340
ccccgaagaa	cgthttccaa	tgatgagcac	ttttaagth	ctgctatgtg	gcgcgggtatt	2400
atcccgtatt	gacgccgggc	aagagcaact	cggtcgccgc	atacactatt	ctcagaatga	2460
cttggttgag	tactcaccag	tcacagaaaa	gcatcttacg	gatggcatga	cagtaagaga	2520
attatgcagt	gctgccataa	ccatgagtga	taacactgcg	gccaacttac	ttctgacaac	2580
gatcggagga	ccgaaggagc	taaccgctth	tttgacaaac	atgggggatc	atgtaactcg	2640
ccttgatcgt	tggaaccgg	agctgaatga	agccatacca	aacgacgagc	gtgacaccac	2700
gatgcctgta	gcaatggcaa	caacgttgcg	caaactatta	actggcgaac	tacttactct	2760
agcttcccgg	caacaattaa	tagactggat	ggaggcggat	aaagttgcag	gaccacttct	2820
gcgctcggcc	cttccggctg	gctggthttat	tgctgataaa	tctggagccg	gtgagcgtgg	2880
gtctcgcggt	atcattgcag	cactggggcc	agatggtaag	ccctcccgt	tcgtagttat	2940
ctacacgacg	gggagtcagg	caactatgga	tgaacgaaat	agacagatcg	ctgagatagg	3000
tgctcactg	attaagcatt	ggtaactgtc	agaccaagth	tactcatata	tactthtagat	3060

tgatttaaaa	cttcattttt	aatttaaaaag	gatctaggtg	aagatccttt	ttgataatct	3120
catgaccaa	atcccttaac	gtgagttttc	gttccactga	gcgtcagacc	ccgtagaaaa	3180
gatcaaagga	tcttcttgag	atcctttttt	tctgcgcgta	atctgctgct	tgcaaacaaa	3240
aaaaccaccg	ctaccagcgg	tggtttgttt	gccggatcaa	gagctaccaa	ctctttttcc	3300
gaaggtaact	ggcttcagca	gagcgcagat	accaaatact	gtcctttctag	tgtagccgta	3360
gttaggccac	cacttcaaga	actctgtagc	accgcctaca	tacctcgctc	tgctaatcct	3420
gttaccagtg	gctgctgcca	gtggcgataa	gtcgtgtctt	accgggttgg	actcaagacg	3480
atagttaccg	gataaggcgc	agcggtcggg	ctgaacgggg	ggttcgtgca	cacagcccag	3540
cttgagcgca	acgacctaca	ccgaactgag	atacctacag	cgtgagctat	gagaaagcgc	3600
cacgcttccc	gaaggggagaa	aggcggacag	gtatccggta	agcggcaggg	tcggaacagg	3660
agagcgcacg	agggagcttc	cagggggaaa	cgcttggtat	ctttatagtc	ctgtcggggt	3720
tcgccacctc	tgacttgagc	gtcgattttt	gtgatgctcg	tcaggggggc	ggagcctatg	3780
gaaaaacgcc	agcaacgcgg	cctttttacg	gttcctggcc	ttttgctggc	cttttgctca	3840
catgttcttt	cctgcgttat	cccctgattc	tgtggataac	cgtattaccg	cctttgagtg	3900
agctgatacc	gctcgccgca	gccgaacgac	cgagcgcagc	gagtcagtga	gcgaggaagc	3960
ggaagagcgc	ccaatacgca	aaccgcctct	ccccgcgcgt	tggccgattc	attaatgcag	4020
ctggcacgac	aggtttcccg	actggaaagc	gggcagtgag	cgcaacgcaa	ttaatgtgag	4080
ttagctcact	cattaggcac	cccaggcttt	acactttatg	cttcgggctc	gtatgtttgtg	4140
tggaattgtg	agcggataac	aatttcacac	aggaaacagc	tatgaccatg	attacgcaa	4200
gcttccaatt	accgtcgctc	gtgatttggt	tgcaaaaaga	acaaaactga	aaaaacccag	4260
acacgctcga	cttctgtctt	tcctattgat	tgcagcttcc	aatttcgtca	cacaacaagg	4320
tcctgtcgac	gcctacttgg	cttcacatac	gttgcatagc	tcgatataga	taataatgat	4380
aatgacagca	ggattatcgt	aatacgtaat	agttgaaaat	ctcaaaaatg	tgtgggtcat	4440
tacgtaaata	atgataggaa	tgggattctt	ctatttttcc	tttttccatt	ctagcagccg	4500
tcgggaaaac	gtggcatcct	ctctttcggg	ctcaattgga	gtcacgctgc	cgtgagcatc	4560
ctctctttcc	atatctaaca	actgagcacg	taaccaatgg	aaaagcatga	gcttagcggt	4620
gctccaaaaa	agtattggat	ggttaatacc	atttgtctgt	tctcttctga	ctttgactcc	4680
tcaaaaaaaaa	aaaatctaca	atcaacagat	cgcttcaatt	acgccctcac	aaaaactttt	4740
ttccttcttc	ttcgcccacg	ttaaatttta	tccctcatgt	tgtctaacgg	atttctgcac	4800
ttgatttatt	ataaaaagac	aaagacataa	tacttctcta	tcaatttcag	ttattgttct	4860
tccttgcggt	attcttctgt	tcttcttttt	cttttgtcat	atataaccat	aaccaagtaa	4920
tacatattca	agtttaaaaca	tgtataccgt	aggacagtac	ttggtagata	gactagaaga	4980
gattggatc	gataaggttt	tcggtgtgcc	aggggattac	aatttgactt	ttctagatta	5040
cattcaaaat	cacgaaggac	tttcttgcca	aggggaatact	aatgaactaa	acgcagcata	5100

20130927_CL5646WOPCT_ST25. txt

tgcagcagat	ggctacgccc	gtgaaagagg	cgtatcagct	cttggttacta	catttcggagt	5160
gggtgaactg	tcagccatta	acggaacagc	tggtagtttt	gcagaacaag	tcctgtcat	5220
ccacatcgtg	ggttctccaa	ctatgaatgt	gcaatccaac	aaaaagctgg	ttcatcattc	5280
cttaggaatg	ggtaactttc	ataactttag	tgaaatggct	aaggaagtca	ctgccgctac	5340
aaccatgctt	actgaagaga	atgcagcttc	agagatcgac	agagtattag	aaacagcctt	5400
gttggaagaa	aggccagtat	acatcaatct	tccaattgat	atagctcata	aagcaatagt	5460
taaacctgca	aaagcactac	aaacagagaa	atcatctggg	gagagagagg	cacaacttgc	5520
agaaatcata	ctatcacact	tagaaaaggc	cgctcaacct	atcgtaatcg	ccgggtcatga	5580
gatcgcccgt	ttccagataa	gagaaagatt	tgaaaactgg	ataaaccaaa	caaagtgtgc	5640
agtaaccaat	ttggcatatg	gcaaaggctc	tttcaatgaa	gagaacgaac	atttcattgg	5700
tacctattac	ccagcttttt	ctgacaaaaa	cgttctggat	tacgttgaca	atagtgaactt	5760
cgttttacat	tttgggtggga	aatcattga	caattctacc	tcctcatttt	ctcaaggctt	5820
taagactgaa	aacactttta	ccgctgcaaa	tgacatcatt	atgctgccag	atgggtctac	5880
ttactctggg	atttctctta	acggtctttt	ggcagagctg	gaaaaactaa	actttacttt	5940
tgctgatact	gctgctaaac	aagctgaatt	agctgttttc	gaaccacagg	ccgaaacacc	6000
actaaagcaa	gacagatttc	accaagctgt	tatgaacttt	ttgcaagctg	atgatgtgtt	6060
ggtcactgag	caggggacat	catctttcgg	tttgatgttg	gcacctctga	aaaaggggat	6120
gaatttgatc	agtcaaacat	tatggggctc	cataggatac	acattacctg	ctatgattgg	6180
ttcacaaatt	gctgccccag	aaaggagaca	cattctatcc	atcgggtgatg	gatcttttca	6240
actgacagca	caggaaatgt	ccaccatctt	cagagagaaa	ttgacaccag	tgatattcat	6300
tatcaataac	gatggctata	cagtcgaaag	agccatccat	ggagaggatg	agagttacaa	6360
tgatatacca	acttgaact	tgcaattagt	tgctgaaaca	tttgggtggg	atgccgaaac	6420
tgtcgacact	cacaacgttt	tcacagaaac	agacttcgct	aatacttttag	ctgctatcga	6480
tgctactcct	caaaaagcac	atgtcgttga	agttcatatg	gaacaaatgg	atatgccaga	6540
atcattgaga	cagattggct	tagccttata	taagcaaaac	tcttaagttt	aaactaagcg	6600
aatttcttat	gatttatgat	ttttattatt	aaataagtta	taaaaaaat	aagtgtatac	6660
aaattttaaa	gtgactctta	ggttttaaaa	cgaaaattct	tattcttgag	taactctttc	6720
ctgtaggtca	ggttgctttc	tcaggtatag	catgaggtcg	ctcttattga	ccacacctct	6780
accggcatgc	cgagcaaagt	cctgcaaata	gctccccatt	tcaccaatt	gtagatatgc	6840
taactccagc	aatgagttga	tgaatctcgg	tgtgtatttt	atgtcctcag	aggacaacac	6900
ctgttgtaat	cgttcttcca	cacg				6924

<210> 187
 <211> 92
 <212> DNA
 <213> Artificial Sequence

```

<220>
<223> LA850 primer

<400> 187
atgactaagc tacactttga cactgctgaa ccagtcaaga tcacacttcc aaatggtttg      60
acataaatta ccgtcgctcg tgatttgttt gc                                     92

<210> 188
<211> 94
<212> DNA
<213> Arti f i c i a l   Sequence

<220>
<223> LA851 primer

<400> 188
ttacaactta attctgacag cttttacttc agtgtatgca tggtagactt cttcacccat      60
ttccaccttg gctaactcgt tgtatcatca ctgg                                     94

<210> 189
<211> 24
<212> DNA
<213> Arti f i c i a l   Sequence

<220>
<223> N1262 primer

<400> 189
cacgtaaggg catgatagaa ttgg                                             24

<210> 190
<211> 26
<212> DNA
<213> Arti f i c i a l   Sequence

<220>
<223> N1263 primer

<400> 190
ggatatagca gttgttgtac actagc                                             26

<210> 191
<211> 6761
<212> DNA
<213> Arti f i c i a l   Sequence

<220>
<223> pLA65 pl asmi d

<400> 191
gatccgcatt gcggattacg tattctaatag ttcagtaccg ttcgtataat gtatgctata      60
cgaagttatg cagattgtac tgagagtgca ccataccacc ttttcaattc atcatttttt      120
ttttattctt ttttttgatt tcggtttcct tgaaattttt ttgattcggg aatctccgaa      180
cagaaggaag aacgaaggaa ggagcacaga cttagattgg tatatatacg catatgtagt      240
gttgaagaaa catgaaattg cccagtattc ttaacccaac tgcacagaac aaaaacctgc      300
aggaaacgaa gataaatcat gtcgaaagct acatataagg aacgtgctgc tactcatcct      360

```

20130927_CL5646WOPCT_ST25. txt

agtcctgttg	ctgccaagct	atttaatatc	atgcacgaaa	agcaaacaaa	cttgtgtgct	420
tcattggatg	ttcgtaccac	caaggaatta	ctggagttag	ttgaagcatt	aggtcccaaa	480
atttgtttac	taaaaacaca	tgtggatatc	ttgactgatt	tttccatgga	gggcacagtt	540
aagccgctaa	aggcattatc	cgccaagtac	aatTTTTTtac	tcttcgaaga	cagaaaattt	600
gctgacattg	gtaatacagt	caaattgcag	tactctgcgg	gtgtatacag	aatagcagaa	660
tgggcagaca	ttacgaatgc	acacggtgtg	gtgggcccag	gtattgttag	cggtttgaag	720
caggcggcag	aagaagtaac	aaaggaacct	agaggccttt	tgatgttagc	agaattgtca	780
tgcaagggct	ccctatctac	tggagaatat	actaagggtta	ctgttgacat	tgcaagagc	840
gacaaagatt	ttgttatcgg	ctttattgct	caaagagaca	tgggtggaag	agatgaaggt	900
tacgattggg	tgattatgac	acccggtgtg	ggtttagatg	acaagggaga	cgcattgggt	960
caacagtata	gaaccgtgga	tgatgtggtc	tctacaggat	ctgacattat	tattgttgga	1020
agaggactat	ttgcaaaggg	aagggatgct	aaggtagagg	gtgaacgtta	cagaaaagca	1080
ggctgggaag	catatttgag	aagatgcggc	cagcaaaact	aaaaaactgt	attataagta	1140
aatgcatgta	tactaaactc	acaaattaga	gcttcaattt	aattatatca	gttattaccc	1200
tatgcggtgt	gaaataccgc	acagatgcgt	aaggagaaaa	taccgcatca	ggaaattgta	1260
aacgttaata	ttttgttaaa	attcgcgtta	aatTTTTgtt	aatcagctc	atTTTTtaac	1320
caataggccg	aaatcggcaa	aatcccttat	aaatcaaaag	aatagaccga	gataggggtg	1380
agtgttgttc	cagtttggaa	caagagtcca	ctattaaaga	acgtggactc	caacgtcaaa	1440
gggcgaaaaa	ccgtctatca	gggcgatggc	ccactacgtg	aaccatcacc	ctaatcaaga	1500
taacttcgta	taatgtatgc	tatacgaacg	gtaccagtga	tgatacaacg	agtttagccaa	1560
ggtgaattca	ctggccgtcg	ttttacaacg	tcgtgactgg	gaaaaccctg	gcgttaccca	1620
acttaatcgc	cttgacgac	atcccccttt	cgccagctgg	cgtaatagcg	aagaggcccg	1680
caccgatcgc	ccttcccaac	agttgcgcag	cctgaatggc	gaatggcgcc	tgatgcggta	1740
ttttctcctt	acgcatctgt	gcggtatttc	acaccgcata	tggtgcactc	tcagtacaat	1800
ctgctctgat	gccgcatagt	taagccagcc	ccgacacccg	ccaacacccg	ctgacgcgcc	1860
ctgacgggct	tgtctgtcc	cggcatccgc	ttacagacaa	gctgtgaccg	tctccgggag	1920
ctgcatgtgt	cagaggtttt	caccgtcatc	accgaaacgc	gcgagacgaa	agggcctcgt	1980
gatacgccta	tttttatagg	ttaatgtcat	gataataatg	gtttcttaga	cgtcagggtg	2040
cacttttcgg	ggaaatgtgc	gcggaacccc	tatttgttta	tttttctaaa	tacattcaaa	2100
tatgtatccg	ctcatgagac	aataaccctg	ataaatgctt	caataatatt	gaaaaaggaa	2160
gagtatgagt	attcaacatt	tccgtgtcgc	ccttattccc	ttttttgcgg	cattttgcct	2220
tcctgttttt	gctcacccag	aaacgctggg	gaaagtaaaa	gatgctgaag	atcagttggg	2280
tgacagagt	ggttacatcg	aactggatct	caacagcggg	aagatccttg	agagttttcg	2340
ccccgaagaa	cgttttccaa	tgatgagcac	ttttaagttt	ctgctatgtg	gcgcgggtatt	2400

20130927_CL5646WOPCT_ST25. txt

atcccgtatt	gacgccgggc	aagagcaact	cggtcgccgc	atacactatt	ctcagaatga	2460
cttggttgag	tactcaccag	tcacagaaaa	gcatcttacg	gatggcatga	cagtaagaga	2520
attatgcagt	gctgccataa	ccatgagtga	taacactgcg	gccaacttac	ttctgacaac	2580
gatcggagga	ccgaaggagc	taaccgcttt	tttgacaaac	atgggggatc	atgtaactcg	2640
ccttgatcgt	tgggaaccgg	agctgaatga	agccatacca	aacgacgagc	gtgacaccac	2700
gatgcctgta	gcaatggcaa	caacgttgcg	caaactatta	actggcgaac	tacttactct	2760
agcttcccgg	caacaattaa	tagactggat	ggaggcggat	aaagttagcag	gaccacttct	2820
gcgctcggcc	cttccggctg	gctggtttat	tgtgataaa	tctggagccg	gtgagcgtgg	2880
gtctcgcggt	atcattgcag	cactggggcc	agatggtaag	ccctcccgtg	tcgtagttat	2940
ctacacgacg	gggagtcagg	caactatgga	tgaacgaaat	agacagatcg	ctgagatagg	3000
tgcctcactg	attaagcatt	ggtaactgtc	agaccaagtt	tactcatata	tacttttagat	3060
tgatttaaaa	cttcattttt	aatttaaaag	gatctaggtg	aagatccttt	ttgataatct	3120
catgacaaa	atcccttaac	gtgagttttc	gttccactga	gcgtcagacc	ccgtagaaaa	3180
gatcaaagga	tcttcttgag	atcctttttt	tctgcgcgta	atctgctgct	tgcaaacaaa	3240
aaaaccaccg	ctaccagcgg	tggtttggtt	gccggatcaa	gagctaccaa	ctctttttcc	3300
gaaggtaact	ggcttcagca	gagcgcagat	accaaatact	gtccttctag	tgtagccgta	3360
gttaggccac	cacttcaaga	actctgtagc	accgcctaca	tacctcgcct	tgctaattct	3420
gttaccagtg	gctgctgcca	gtggcgataa	gtcgtgtctt	accgggttgg	actcaagacg	3480
atagttaccg	gataaggcgc	agcggtcggg	ctgaacgggg	ggttcgtgca	cacagcccag	3540
cttggagcga	acgacctaca	ccgaactgag	atacctacag	cgtgagctat	gagaaagcgc	3600
cacgcttccc	gaaggggaga	aggcggacag	gtatccggta	agcggcaggg	tcggaacagg	3660
agagcgcacg	agggagcttc	cagggggaaa	cgcctggtat	ctttatagtc	ctgtcggggt	3720
tcgccacctc	tgacttgagc	gtcgattttt	gtgatgctcg	tcaggggggc	ggagcctatg	3780
gaaaaacgcc	agcaacgcgg	ccttttttacg	gttcctggcc	ttttgctggc	cttttgctca	3840
catgttcttt	cctgcgttat	cccctgattc	tgtggataac	cgtattaccg	cctttgagtg	3900
agctgatacc	gctcgccgca	gccgaacgac	cgagcgcagc	gagtcagtga	gcgaggaagc	3960
ggaagagcgc	ccaatacgca	aaccgcctct	ccccgcgcgt	tggccgattc	attaatgcag	4020
ctggcacgac	aggtttcccg	actggaaagc	gggcagtgag	cgcaacgcaa	ttaatgtgag	4080
ttagctcact	cattaggcac	cccaggcttt	acactttatg	cttccggctc	gtatgttgtg	4140
tggatttggt	agcggataac	aatttcacac	aggaaacagc	tatgaccatg	attacgcaa	4200
gcttacctgg	taaaacctct	agtggagtag	tagatgtaat	caatgaagcg	gaagccaaaa	4260
gaccagagta	gaggcctata	gaagaaactg	cgataccttt	tgtgatggct	aaacaaacag	4320
acatcttttt	atatgttttt	acttctgtat	atcgtgaagt	agtaagtgat	aagcgaattt	4380
ggctaagaac	gttgtaagtg	aacaagggac	ctcttttgcc	tttcaaaaaa	ggattaaatg	4440

20130927_CL5646WOPCT_ST25. txt

gagttaatca	ttgagattta	gttttcgtta	gattctgtat	ccctaaataa	ctcccttacc	4500
cgacgggaag	gcacaaaaga	cttgaataat	agcaaacggc	cagtagccaa	gaccaaataa	4560
tactagagtt	aactgatggg	cttaaacagg	cattacgtgg	tgaactccaa	gaccaatata	4620
caaaatatcg	ataagttatt	cttgcccacc	aatttaagga	gcctacatca	ggacagtagt	4680
accattcctc	agagaagagg	tatacataac	aagaaaatcg	cgtgaacacc	ttatataact	4740
tagcccgtta	ttgagctaaa	aaaccttgca	aaatttccta	tgaataagaa	tacttcagac	4800
gtgataaaaa	tttactttct	aactcttctc	acgctgcccc	tatctgttct	tccgctctac	4860
cgtgagaaat	aaagcatcga	gtacggcagt	tcgctgtcac	tgaactaaaa	caataaggct	4920
agttcgaatg	atgaacttgc	ttgctgtcaa	acttctgagt	tgccgctgat	gtgacactgt	4980
gacaataaat	tcaaaccggg	tatagcggtc	tcctccggta	ccggttctgc	cacctccaat	5040
agagctcagt	aggagtcaga	acctctgcgg	tggctgtcag	tgactcatcc	gcgtttcgt	5100
agttgtgctc	gtgcacattt	cgcccgttcc	cgctcatctt	gcagcaggcg	gaaattttca	5160
tcacgctgta	ggacgcaaaa	aaaaaataat	taatcgtaca	agaatcttgg	aaaaaaaatt	5220
gaaaaatttt	gtataaaagg	gatgacctaa	cttgactcaa	tggctttttac	accaggtatt	5280
ttccctttcc	ttgtttgtta	caattataga	agcaagacaa	aaacatatag	acaacctatt	5340
cctaggagtt	atattttttt	accctaccag	caatataagt	aaaaaactgt	ttatgaaagc	5400
attagtgtat	agggggccag	gccagaagtt	ggtggaagag	agacagaagc	cagagcttaa	5460
ggaacctggg	gacgctatag	tgaaggtaac	aaagactaca	atttgcgga	ccgatctaca	5520
cattcttaaa	ggtgacgttg	cgacttgtaa	acccggctgt	gtattagggc	atgaaggagt	5580
gggggttatt	gaatcagtcg	gatctggggg	tactgctttc	caaccaggcg	atagagtttt	5640
gatatcatgt	atatcgagtt	gcggaaagtg	ctcattttgt	agaagaggaa	tgttcagtca	5700
ctgtacgacc	gggggttggg	ttctgggcaa	cgaaattgat	ggtaccaag	cagagtacgt	5760
aagagtacca	catgctgaca	catcccttta	tcgtattccg	gcagggtcgg	atgaagaggc	5820
cttagtcatg	ttatcagata	ttctaccaac	gggttttgag	tgcgaggtcc	taaacggcaa	5880
agtcgcacct	ggttcttcgg	tggctatagt	aggtgctggg	cccgttgggt	tggccgcctt	5940
actgacagca	caattctact	ccccagctga	aatcataatg	atcgatcttg	atgataacag	6000
gctgggatta	gccaaacaat	ttggtgccac	cagaacagta	aactccacgg	gtggtaacgc	6060
cgagccgaa	gtgaaagctc	ttactgaagg	cttaggtgtt	gatactgcca	ttgaagcagt	6120
tgggatacct	gctacatttg	aattgtgtca	gaatatcgta	gctcccgggtg	gaactatcgc	6180
taatgtcggc	gttcacggta	gcaaagttga	tttgcatctt	gaaagtttat	ggtcccataa	6240
tgtcacgatt	actacaagg	tggttgacac	ggctaccacc	ccgatgttac	tgaaaactgt	6300
tcaaagtcac	aagctagatc	catctagatt	gataacacat	agattcagcc	tggaccagat	6360
cttggacgca	tatgaaactt	ttggccaagc	tgctgtctact	caagcactaa	aagtcatcat	6420
ttcgatggag	gcttgattaa	ttaagagtaa	gcgaatttct	tatgatttat	gattttttatt	6480

attaaataag ttataaaaaa aataagtgtgta tacaaatttt aaagtgactc ttaggtttta 6540
 aaacgaaaat tcttattctt gagtaactct ttctgttagg tcaggttgct ttctcaggta 6600
 tagcatgagg tcgctcttat tgaccacacc tctaccggca tgccgagcaa atgcctgcaa 6660
 atcgctcccc atttcacca attgtagata tgctaactcc agcaatgagt tgatgaatct 6720
 cgggtgtgtat tttatgtcct cagaggacaa cacctgtggg g 6761

<210> 192
 <211> 80
 <212> DNA
 <213> Arti f i c i a l Sequence

<220>
 <223> LA855 primer

<400> 192
 gcacaatatt tcaagctata ccaagcatat aatcaactat ctcatatata acctggtaaa 60
 acctctagtg gagtagtaga 80

<210> 193
 <211> 83
 <212> DNA
 <213> Arti f i c i a l Sequence

<220>
 <223> LA856 primer

<400> 193
 gcttatttag aagtgtcaac aacgtatcta ccaacgattt gacccttttc cacaccttgg 60
 ctaactcggt gtatcatcac tgg 83

<210> 194
 <211> 25
 <212> DNA
 <213> Arti f i c i a l Sequence

<220>
 <223> LA414 primer

<400> 194
 ccagagctga tgaggggtat ctcga 25

<210> 195
 <211> 25
 <212> DNA
 <213> Arti f i c i a l Sequence

<220>
 <223> LA749 primer

<400> 195
 caagtctttt gtgccttccc gtcgg 25

<210> 196
 <211> 25
 <212> DNA
 <213> Arti f i c i a l Sequence

<220>
 <223> LA413 primer

 <400> 196
 ggacataaaa tacacaccga gattc 25

 <210> 197
 <211> 90
 <212> DNA
 <213> Arti f i c i a l Sequence

 <220>
 <223> LA860 primer

 <400> 197
 tctcaattat tattttctac tcataacctc acgcaaaata acacagtcaa atcaatcaaa 60
 atgaaagcat tagtgtatag gggcccaggc 90

 <210> 198
 <211> 26
 <212> DNA
 <213> Arti f i c i a l Sequence

 <220>
 <223> N1093 primer

 <400> 198
 tttcaagatg caaatcaact ttgcta 26

 <210> 199
 <211> 20
 <212> DNA
 <213> Arti f i c i a l Sequence

 <220>
 <223> LA681 primer

 <400> 199
 ttattgctta gcgttggtag 20

 <210> 200
 <211> 9612
 <212> DNA
 <213> Arti f i c i a l Sequence

 <220>
 <223> Pl asmi d pHR81-K9D3

 <400> 200
 aaacagtatg gaagaatgta agatggctaa gatttactac caagaagact gtaacttgtc 60
 cttgttggat ggtaagacta tcgccgttat cggttacggt tctcaaggtc acgctcatgc 120
 cctgaatgct aaggaatccg gttgtaacgt tatcattggt ttatacgaag gtgctaagga 180
 ttggaaaaga gctgaagaac aagggtttcga agtctacacc gctgctgaag ctgctaagaa 240
 ggctgacatc attatgatct tgatcaacga tgaaaagcag gctaccatgt acaaaaacga 300
 catcgaacca aacttgaag ccggtaacat gttgatgttc gtcacggtt tcaacatcca 360
 tttcggttgt attgttcac caaaggacgt tgatgtcact atgatcgctc caaaggggtcc 420

20130927_CL5646WOPCT_ST25. txt

aggtcacacc	gtagatccg	aatacgaaga	aggtaaaggt	gtcccatgct	tggttgctgt	480
cgaacaagac	gctactggca	aggcttttga	tatggctttg	gcctacgctt	tagccatcgg	540
tggtgctaga	gccggtgtct	tggaaactac	cttcagaacc	gaaactgaaa	ccgacttggt	600
cggtgaacaa	gctgttttat	gtggtggtgt	ctgcgctttg	atgcaggccg	gttttgaaac	660
cttggttgaa	gccggttacg	acccaagaaa	cgcttacttc	gaatgtatcc	acgaaatgaa	720
gttgatcggt	gacttgatct	accaatctgg	tttctccggt	atgcgttact	ctatctccaa	780
cactgctgaa	tacggtgact	acattaccgg	tccaaagatc	attactgaag	ataccaagaa	840
ggctatgaag	aagattttgt	ctgacattca	agatggtacc	tttgccaagg	acttcttggt	900
tgacatgtct	gatgctgggt	cccagggtcca	cttcaaggct	atgagaaagt	tggcctccga	960
acaccagct	gaagttgtcg	gtgaagaaat	tagatccttg	tactcctggt	ccgacgaaga	1020
caagttgatt	aacaactgat	attttctctt	ggccctgcag	gcctatcaag	tgctggaaac	1080
tttttctctt	ggaatttttg	caacatcaag	tcatagtcaa	ttgaattgac	ccaatttcac	1140
atttaagatt	tttttttttt	catccgacat	acatctgtac	actaggaagc	cctgtttttc	1200
tgaagcagct	tcaaataatat	atatttttta	catattttatt	atgattcaat	gaacaatcta	1260
attaaatcga	aaacaagaac	cgaaacgcga	ataaataatt	tatttagatg	gtgacaagtg	1320
tataagtcct	catcgggaca	gctacgattt	ctctttcggt	tttggtgag	ctactggttg	1380
ctgtgacgca	gcggcattag	cgcggcgtta	tgagctaccc	tcgtggcctg	aaagatggcg	1440
ggaataaagc	ggaactaaaa	attactgact	gagccatatt	gaggtcaatt	tgtcaactcg	1500
tcaagtcacg	tttggtggac	ggcccctttc	caacgaatcg	tatatactaa	catgcgcgcg	1560
cttcctatat	acacatatat	atatatatat	atatatatat	gtgtgcgtgt	atgtgtacac	1620
ctgtatttaa	tttcttact	cgcggttttt	tcttttttct	caattcttgg	cttcctcttt	1680
ctcgagcggg	ccggtatcct	cgcggtgccc	gcagatctat	ttaaatggcg	cgccgacgtc	1740
aggtggcact	tttcggggaa	atgtgcgcgg	aacccttatt	tgtttatttt	tctaaataca	1800
ttcaaatatg	tatccgctca	tgagacaata	accctgataa	atgcttcaat	aatattgaaa	1860
aaggaagagt	atgagtattc	aacatttccg	tgtcgccctt	attccctttt	ttgcggcatt	1920
ttgccttcct	gtttttgctc	accagaaaac	gctggtgaaa	gtaaaagatg	ctgaagatca	1980
gttgggtgca	cgagtgggtt	acatcgaact	ggatctcaac	agcggtaaga	tccttgagag	2040
ttttcgcccc	gaagaacgtt	ttccaatgat	gagcactttt	aaagtctctg	tatgtggcgc	2100
ggtattatcc	cgtattgacg	ccgggcaaga	gcaactcggg	cgccgcatac	actattctca	2160
gaatgacttg	gttgagtact	caccagtcac	agaaaagcat	cttacggatg	gcatgacagt	2220
aagagaatta	tgcagtgtct	ccataacat	gagtataaac	actgcggcca	acttacttct	2280
gacaacgata	ggaggaccga	aggagctaac	cgcttttttg	cacaacatgg	gggatcatgt	2340
aactcgcctt	gatcgttggg	aaccggagct	gaatgaagcc	ataccaaacg	acgagcgtga	2400
caccacgatg	cctgtagcaa	tggcaacaac	gttgcgcaaa	ctattaactg	gcgaactact	2460

20130927_CL5646WOPCT_ST25.txt

tactctagct	tcccggcaac	aattaataga	ctggatggag	gcgataaag	ttgcaggacc	2520
acttctgcgc	tcggcccttc	cggctggctg	gtttattgct	gataaatctg	gagccggtga	2580
gcgtgggtct	cgcggtatca	ttgcagcact	ggggccagat	ggtaagccct	cccgtatcgt	2640
agttatctac	acgacgggga	gtcaggcaac	tatggatgaa	cgaaatagac	agatcgctga	2700
gatagggtgcc	tcactgatta	agcattggta	actgtcagac	caagtttact	catatatact	2760
ttagattgat	ttaaaacttc	atttttaatt	taaaaggatc	taggtgaaga	tcctttttga	2820
taatctcatg	acaaaaatcc	cttaacgtga	gttttcgttc	cactgagcgt	cagaccccgt	2880
agaaaagatc	aaaggatctt	cttgagatcc	tttttttctg	cgcgtaatct	gctgcttgca	2940
aacaaaaaaaa	ccaccgctac	cagcggtggt	ttgtttgccg	gatcaagagc	taccaactct	3000
ttttccgaag	gtaactggct	tcagcagagc	gcagatacca	aatactgttc	ttctagtgtg	3060
gccgtagtta	ggccaccact	tcaagaactc	tgtagcaccg	cctacatacc	tcgctctgct	3120
aatcctgtta	ccagtggctg	ctgccagtgg	cgataagtcg	tgtcttaccg	ggttggactc	3180
aagacgatag	ttaccggata	aggcgcagcg	gtcgggctga	acgggggggtt	cgtgcacaca	3240
gccagcttg	gagcgaacga	cctacaccga	actgagatac	ctacagcgtg	agctatgaga	3300
aagcgccacg	cttcccgaag	ggagaaaggc	ggacaggtat	ccggtaagcg	gcagggtcgg	3360
aacaggagag	cgcacgaggg	agcttcagg	gggaaacgcc	tggatatctt	atagtcctgt	3420
cgggtttcgc	cacctctgac	ttgagcgtcg	atttttgtga	tgctcgtcag	gggggcggag	3480
cctatggaaa	aacgccagca	acgggcctt	tttacggttc	ctggcctttt	gctggccttt	3540
tgctcacatg	ttctttcctg	cgttatcccc	tgattctgtg	gataaccgta	ttaccgcctt	3600
tgagtgagct	gataccgctc	gccgcagccg	aacgaccgag	cgcagcgagt	cagtgagcga	3660
ggaagcggaa	gagcgcccaa	tacgcaaacc	gcctctcccc	gcgcgttggc	cgattcatta	3720
atgcagctgg	cacgacaggt	ttcccgaactg	gaaagcgggc	agtgagcgca	acgcaattaa	3780
tgtgagttag	ctcactcatt	aggcacccca	ggctttacac	tttatgcttc	cggctcgtat	3840
gttgtgtgga	attgtgagcg	gataacaatt	tcacacagga	aacagctatg	accatgatta	3900
cgccaagctt	tttctttcca	attttttttt	tttcgtcatt	ataaaaaatca	ttacgaccga	3960
gattcccggg	taataactga	tataattaa	ttgaagctct	aattttgtgag	tttagtatac	4020
atgcatttac	ttataatata	gttttttagt	tttgctggcc	gcattcttctc	aaatatgctt	4080
cccagcctgc	ttttctgtaa	cgttcaccct	ctaccttagc	atcccttccc	tttgcaaata	4140
gtcctcttcc	aacaataata	atgtcagatc	ctgtagagac	cacatcatcc	acggttctat	4200
actgttgacc	caatgcgtct	cccttgtcat	ctaaaccac	accgggtgtc	ataatcaacc	4260
aatcgtaacc	ttcatctctt	ccacccatgt	ctctttgagc	aataaagccg	ataacaaaat	4320
ctttgtcgct	cttcgcaatg	tcaacagtac	ccttagtata	ttctccagta	gatagggagc	4380
ccttgcatga	caattctgct	aacatcaaaa	ggcctctagg	ttcctttgtt	acttcttctg	4440
ccgcctgctt	caaaccgcta	acaatacctg	ggcccaccac	accgtgtgca	ttcgtaatgt	4500

20130927_CL5646WOPCT_ST25. txt

ctgcccattc	tgctattctg	tatacacccg	cagagtactg	caatttgact	gtattaccaa	4560
tgtcagcaaa	ttttctgtct	tcgaagagta	aaaaattgta	cttggcggat	aatgccttta	4620
gcggcctaac	tgtgccctcc	atggaaaaat	cagtcaagat	atccacatgt	gttttttagta	4680
aacaaatfff	gggacctaaf	gcttcaacta	actccagtaa	ttccttggtg	gtacgaacat	4740
ccaatgaagc	acacaagfff	gtttgctfff	cgtgcatgat	attaaatagc	ttggcagcaa	4800
caggactagg	atgagtagca	gcacgttcct	tatatgtagc	tttcgacatg	atttatcttc	4860
gtttcctgca	ggtttttggt	ctgtgcagtt	gggttaagaa	tactgggcaa	tttcatgttt	4920
cttcaacact	acatatgcgt	atatatacca	atctaagtct	gtgctccttc	cttcgttctt	4980
ccttctgttc	ggagattacc	gaatcaaaaa	aatttcaagg	aaaccgaaat	caaaaaaaag	5040
aataaaaaaa	aatgatgaa	ttgaaaagct	tgcatgcctg	caggtcgact	ctagtatact	5100
ccgtctactg	tacgatacac	ttccgctcag	gtccttgtcc	tttaacgagg	ccttaccact	5160
cttttgttac	tctattgatc	cagctcagca	aaggcagtgt	gatctaagat	tctatcttcg	5220
cgatgtagta	aaactagcta	gaccgagaaa	gagactagaa	atgcaaaagg	cacttctaca	5280
atggctgcca	tcattattat	ccgatgtgac	gctgcatttt	tttttttttt	tttttttttt	5340
tttttttttt	tttttttttt	ttttttttgt	acaatatatca	taaaaaaaga	gaatcttttt	5400
aagcaaggat	tttcttaact	tcttcggcga	cagcatcacc	gacttcggtg	gtactgttgg	5460
aaccacctaa	atcaccagtt	ctgatacctg	catccaaaac	ctttttaact	gcatcttcaa	5520
tggctttacc	ttcttcaggc	aagttcaatg	acaatttcaa	catcattgca	gcagacaaga	5580
tagtggcgat	agggttgacc	ttattctttg	gcaaactctgg	agcggaacca	tggcatgggt	5640
cgtacaaacc	aatgcggtg	ttcttgctctg	gcaaagaggc	caaggacgca	gatggcaaca	5700
aacccaagga	gcctgggata	acggaggctt	catcggagat	gatatcacca	aacatgtttgc	5760
tggtgattat	aataccattt	aggtgggttg	ggttcttaac	taggatcatg	gcggcagaat	5820
caatcaattg	atgttgaact	ttcaatgtag	ggaattcggt	cttgatgggt	tcctccacag	5880
tttttctcca	taatcttgaa	gaggccaaaa	cattagcttt	atccaaggac	caaataggca	5940
atggtggctc	atgtttagg	gccatgaaag	cggccattct	tgtgattctt	tgcacttctg	6000
gaacggtgta	ttgttacta	tccaagcgga	caccatcacc	atcgtcttcc	tttctcttac	6060
caaagtaaaf	acctcccact	aattctctaa	caacaacgaa	gtcagtacct	ttagcaaatt	6120
gtggcttgat	tggagataag	tctaaaagag	agtcggatgc	aaagttacat	ggtcttaagt	6180
tggcgtacaa	ttgaagttct	ttacggattt	ttagtaaacc	ttgttcagggt	ctaacactac	6240
cggtagccca	tttaggacca	cccacagcac	ctaacaaaac	ggcatcagcc	ttcttgaggg	6300
cttccagcgc	ctcatctgga	agtggaacac	ctgtagcatc	gatagcagca	ccaccaatta	6360
aatgattttc	gaaatcgaac	ttgacattgg	aacgaacatc	agaaatagct	ttaagaacct	6420
taatggcttc	ggctgtgatt	tcttgaccaa	cgtggtcacc	tggcaaaacg	acgatcttct	6480
taggggcaga	cattacaatg	gtatatcctt	gaaatatata	taaaaaaaaa	aaaaaaaaaa	6540

20130927_CL5646WOPCT_ST25. txt

aaaaaaaaa	atgcagcttc	tcaatgatat	tcgaatacgc	tttgaggaga	tacagcctaa	6600
tatccgacaa	actgttttac	agattttacga	tcgtacttgt	tacccatcat	tgaattttga	6660
acatccgaac	ctgggagttt	tccctgaaac	agatagtata	tttgaacctg	tataataata	6720
tatagtctag	cgctttacgg	aagacaatgt	atgtatttcg	gttcctggag	aaactattgc	6780
atctattgca	taggtaatct	tgcacgtcgc	atccccggtt	cattttctgc	gtttccatct	6840
tgcacttcaa	tagcatatct	ttgttaacga	agcatctgtg	cttcattttg	tagaacaaaa	6900
atgcaacgcg	agagcgctaa	tttttcaaac	aaagaatctg	agctgcattt	ttacagaaca	6960
gaaatgcaac	gcgaaagcgc	tattttacca	acgaagaatc	tgtgcttcat	ttttgtaaaa	7020
caaaaatgca	acgcgagagc	gctaattttt	caaacaaaga	atctgagctg	cattttttaca	7080
gaacagaaat	gcaacgcgag	agcgctattt	taccaacaaa	gaatctatac	ttcttttttg	7140
ttctacaaaa	atgcatcccg	agagcgctat	ttttctaaca	aagcatctta	gattactttt	7200
tttctccttt	gtgcgctcta	taatgcagtc	tcttgataac	tttttgact	gtaggtccgt	7260
taaggttaga	agaaggctac	tttgggtgtct	attttctctt	ccataaaaaa	agcctgactc	7320
cacttccgcg	gtttactgat	tactagcgaa	gctgcgggtg	cattttttca	agataaaggc	7380
atccccgatt	atattctata	ccgatgtgga	ttgcgcatac	tttgtgaaca	gaaagtgata	7440
gcgttgatga	ttcttcattg	gtcagaaaaat	tatgaacggt	ttcttctatt	ttgtctctat	7500
atactacgta	taggaaatgt	ttacattttc	gtattgtttt	cgattcactc	tatgaatagt	7560
tcttactaca	atttttttgt	ctaaagagta	atactagaga	taaacataaa	aatgttagag	7620
gtcgagttta	gatgcaagtt	caaggagcga	aagggtgatg	ggtaggttat	atagggatat	7680
agcacagaga	tatatagcaa	agagatactt	ttgagcaatg	tttgtggaag	cggtattcgc	7740
aatattttag	tagctcgtta	cagtccggtg	cgtttttgggt	tttttgaaag	tgcgtcttca	7800
gagcgctttt	ggttttcaaa	agcgctctga	agttcctata	ctttctagag	aataggaact	7860
tcggaatagg	aacttcaaag	cgtttccgaa	aacgagcgct	tccgaaaatg	caacgcgagc	7920
tgcgcacata	cagctcactg	ttcacgtcgc	acctatatct	gcgtgttgcc	tgtatatata	7980
tatacatgag	aagaacggca	tagtgcggtg	ttatgcttaa	atgcgtactt	atatgcgtct	8040
atttatgtag	gatgaaaggt	agtctagtac	ctcctgtgat	attatcccat	tccatgcggg	8100
gtatcgtatg	cttccttcag	cactaccctt	tagctgttct	atatgctgcc	actcctcaat	8160
tggattagtc	tcatccttca	atgctatcat	ttcctttgat	attggatcat	atgcatagta	8220
ccgagaaact	agaggatctc	ccattaccga	catttgggcg	ctatacgtgc	atatgttcat	8280
gtatgtatct	gtatttaaaa	cacttttgta	ttatttttcc	tcatatatgt	gtataggttt	8340
atacggatga	tttaattatt	acttcaccac	cctttatttc	aggctgatat	cttagccttg	8400
ttactagtca	ccggtggcgg	ccgcacctgg	taaaacctct	agtggagtag	tagatgtaat	8460
caatgaagcg	gaagccaaaa	gaccagagta	gaggcctata	gaagaaactg	cgataccttt	8520
tgtgatggct	aaacaaacag	acatcttttt	atatgttttt	acttctgtat	atcgtgaagt	8580

20130927_CL5646W0PCT_ST25. txt

agtaagtgat	aagcgaattt	ggctaagaac	gttgtaagtg	aacaagggac	ctcttttgcc	8640
tttcaaaaaa	ggattaaatg	gagttaatca	ttgagattta	gttttcgtta	gattctgtat	8700
ccctaaataa	ctcccttacc	cgacgggaag	gcacaaaaga	cttgaataat	agcaaacggc	8760
cagtagccaa	gaccaaataa	tactagagtt	aactgatggg	cttaaacagg	cattacgtgg	8820
tgaactccaa	gaccaatata	caaaatatcg	ataagttatt	cttgcccacc	aatttaagga	8880
gcctacatca	ggacagtagt	accattcctc	agagaagagg	tatacataac	aagaaaatcg	8940
cgpgaacacc	ttatataact	tagcccgtta	ttgagctaaa	aaaccttgca	aaatttccta	9000
tgaataagaa	tacttcagac	gtgataaaaa	tttactttct	aactcttctc	acgctgcccc	9060
tatctgttct	tccgctctac	cgtgagaaat	aaagcatcga	gtacggcagt	tcgctgtcac	9120
tgaactaaaa	caataaggct	agttcgaatg	atgaacttgc	ttgctgtcaa	acttctgagt	9180
tgccgctgat	gtgacactgt	gacaataaat	tcaaaccggg	tatagcgggc	tcctccggta	9240
ccggttctgc	cacctccaat	agagctcagt	aggagtcaga	acctctgcgg	tggctgtcag	9300
tgactcatcc	gcgtttcgta	agttgtgcgc	gtgcacattt	cgcccgttcc	cgctcatctt	9360
gcagcaggcg	gaaattttca	tcacgctgta	ggacgcaaaa	aaaaaataat	taatcgta	9420
agaatcttgg	aaaaaaaaatt	gaaaaatttt	gtataaaagg	gatgacctaa	cttgactcaa	9480
tggcttttac	accaggtatt	ttccctttcc	ttgtttgtta	caattataga	agcaagacaa	9540
aaacatatag	acaacctatt	cctaggagtt	atattttttt	accctaccag	caatataagt	9600
aaaaaactgt	tt					9612

<210>	201
<211>	7938
<212>	DNA
<213>	Artificial Sequence

<220>
<223> PI asmi d pYZ067Dki vDDadh

[illegible]

20130927_CL5646WOPCT_ST25. txt

attaacgtcc	acacaggtat	agggtttctg	gaccatatga	tacatgctct	ggccaagcat	720
tccggctggt	cgctaactgt	tgagtgcatt	ggtgacttac	acatagacga	ccatcacacc	780
actgaagact	gcgggattgc	tctcgggtcaa	gcttttaaag	aggccctagg	ggccgtgcgt	840
ggagtaaaaa	ggtttggatc	aggatttgcg	cctttggatg	aggcactttc	cagagcgggtg	900
gtagatcttt	cgaacaggcc	gtacgcagtt	gtcgaacttg	gtttgcaaag	ggagaaagta	960
ggagatctct	cttgcgagat	gatcccgcat	tttcttgaaa	gctttgcaga	ggctagcaga	1020
attaccctcc	acgttgattg	tctgcgaggc	aagaatgata	atcaccgtag	tgagagtgcg	1080
ttcaaggctc	ttgcggttgc	cataagagaa	gccacctcgc	ccaatggtac	caacgatgtt	1140
ccctccacca	aagggtgttct	tatgtagtga	caccgattat	ttaaagctgc	agcatatcat	1200
atatatacat	gtgtatatat	gtatacctat	gaatgtcagt	aagtatgtat	acgaacagta	1260
tgatactgaa	gatgacaagg	taatgcatca	ttctatacgt	gtcatttctga	acgaggcgcg	1320
ctttcctttt	ttctttttgc	tttttctttt	tttttctctt	gaactcgacg	gatctatgcg	1380
gtgtgaaata	ccgcacagat	gcgtaaggag	aaaataccgc	atcaggaaat	tgtaagcggt	1440
aatattttgt	taaaattcgc	gttaaatttt	tgtaaataca	gtcatttttt	taaccaatag	1500
gccgaaatcg	gcaaaatccc	ttataaatca	aaagaataga	ccgagatagg	gttgagtgtt	1560
gttccagttt	ggaacaagag	tccactatta	aagaacgtgg	actccaacgt	caaagggcga	1620
aaaaccgtct	atcagggcga	tggcccacta	cgtggccggc	ttcacatacg	ttgcatacgt	1680
cgatatagat	aataatgata	atgacagcag	gattatcgta	atacgtataa	gctgaaaatc	1740
tcaaaaatgt	gtgggtcatt	acgtaaataa	tgataggaat	gggattcttc	tatttttctt	1800
ttttccattc	tagcagccgt	cgggaaaacg	tggcatcctc	tctttcgggc	tcaattggag	1860
tcacgctgcc	gtgagcatcc	tctctttcca	tatctaacaa	ctgagcacgt	aaccaatgga	1920
aaagcatgag	cttagcgttg	ctccaaaaaa	gtattggatg	gttaatacca	tttgtctgtt	1980
ctcttctgac	tttgactcct	caaaaaaaaa	aatctacaat	caacagatcg	cttcaattac	2040
gccctcacia	aaactttttt	ccttcttctt	cgcccacgtt	aaattttatc	cctcatgttg	2100
tctaacggat	ttctgcactt	gatttattat	aaaaagacaa	agacataata	cttctctatc	2160
aatttcagtt	attgttcttc	cttgcgttat	tcttctgttc	ttctttttct	tttgtcatat	2220
ataaccataa	ccaagtaata	catattcaaa	cacgtgagta	tgactgacaa	aaaaactctt	2280
aaagacttaa	gaaatcgtag	ttctgtttac	gattcaatgg	ttaaatacacc	taatcgtgct	2340
atgttgcgtg	caactgggtat	gcaagatgaa	gactttgaaa	aacctatcgt	cggtgtcatt	2400
tcaacttggg	ctgaaaacac	accttgtaat	atccacttac	atgactttgg	taaactagcc	2460
aaagtgcgtg	ttaaggaagc	tgggtgcttg	ccagttcagt	tcggaacaat	cacggtttct	2520
gatggaatcg	ccatgggaac	ccaaggaatg	cgtttctcct	tgacatctcg	tgatattatt	2580
gcagattcta	ttgaagcagc	catgggaggt	cataatgcgg	atgcttttgt	agccattggc	2640
ggttgtgata	aaaacatgcc	cggttctgtt	atcgctatgg	ctaacaatgga	tatcccagcc	2700

20130927_CL5646WOPCT_ST25. txt

atTTTTgctt acggcggaac aattgcacct ggtaatttag acggcaaaga tatcgattta	2760
gtctctgtct ttgaagggtgt cggccattgg aaccacggcg atatgaccaa agaagaagtt	2820
aaagctttgg aatgtaatgc ttgtcccgggt cctggaggct gcggtggtat gtatactgct	2880
aacacaatgg cgacagctat tgaagttttg ggacttagcc ttccgggttc atcttctcac	2940
ccggctgaat ccgcagaaaa gaaagcagat attgaagaag ctggtcgcgc tgttgtcaaa	3000
atgctcgaaa tgggcttaaa accttctgac attttaacgc gtgaagcttt tgaagatgct	3060
attactgtaa ctatggctct gggagggttca accaactcaa cccttcacct cttagctatt	3120
gcccatgctg ctaatgtgga attgacactt gatgatttca atactttcca agaaaaagtt	3180
cctcatttgg ctgatttgaa accttctggt caatatgtat tccaagacct ttacaaggctc	3240
ggaggggtac cagcagttat gaaatatctc cttaaaaatg gcttccttca tggtgaccgt	3300
atcacttgta ctggcaaaac agtcgctgaa aatttgaagg cttttgatga tttaacacct	3360
ggtcaaaagg ttattatgcc gcttgaaaat cctaaacgtg aagatgggtcc gctcattatt	3420
ctccatggta acttggctcc agacggtgcc gttgccaaag tttctggtgt aaaagtgcgt	3480
cgtcatgtcg gtctgtctaa ggtctttaat tctgaagaag aagccattga agctgtcttg	3540
aatgatgata ttgttgatgg tgatgttgtt gtcgtacgtt ttgtaggacc aaagggcgggt	3600
cctggtatgc ctgaaatgct ttccctttca tcaatgattg ttggtaaagg gcaaggtgaa	3660
aaagttgccc ttctgacaga tggccgcttc tcagggtggt cttatggtct tgtcgtgggt	3720
catatcgctc ctgaagcaca agatggcgggt ccaatcgctt acctgcaaac aggagacata	3780
gtcactattg accaagacac taaggaatta cactttgata tctccgatga agagttaaaa	3840
catcgtcaag agaccattga attgccaccg ctctattcac gcggtatcct tggtaaatat	3900
gctcacatcg tttcgtctgc ttctagggga gccgtaacag acttttggaa gcctgaagaa	3960
actggcaaaa aatgttgtcc tggttgctgt ggttaagcgg ccgcgttaat tcaaattaat	4020
tgatatagtt ttttaatgag tattgaatct gtttagaaat aatggaatat tatttttatt	4080
tatttattta tattattggt cggctctttt cttctgaagg tcaatgacaa aatgatatga	4140
aggaaataat gatttctaaa attttacaac gtaagatatt tttaaaaaag cctagctcat	4200
cttttgtcat gcactatttt actcacgctt gaaattaacg gccagtcacac tgcggagtca	4260
tttcaaagtc atcctaatcg atctatcgtt tttgatagct cattttggag ttcgcgagga	4320
tcccagcttt tgttcccttt agtgagggtt aattgcgcgc ttggcgtaat catggtcata	4380
gctgtttcct gtgtgaaatt gttatccgct cacaattcca cacaacatac gagccggaag	4440
cataaagtgt aaagcctggg gtgcctaata agtgagctaa ctcacattaa ttgcgttgcg	4500
ctcactgccc gctttccagt cgggaaacct gtcgtgccag ctgcattaat gaatcggcca	4560
acgcgcgggg agaggcggtt tgcgtattgg gcgctcttcc gcttcctcgc tcaactgactc	4620
gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg cggtaatcgc	4680
gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaaag gccagcaaaa	4740

20130927_CL5646WOPCT_ST25. txt

ggccaggaac	cgtaaaaagg	ccgcgttgct	ggcgtttttc	cataggctcc	gccccctga	4800
cgagcatcac	aaaaatcgac	gctcaagtca	gaggtggcga	aacccgacag	gactataaag	4860
ataccaggcg	tttccccctg	gaagctccct	cgtgcgctct	cctgttccga	ccctgccgct	4920
taccggatac	ctgtccgcct	ttctcccttc	gggaagcgtg	gcgctttctc	atagctcacg	4980
ctgtaggtat	ctcagttcgg	tgtaggtcgt	tcgctccaag	ctgggctgtg	tgcacgaacc	5040
ccccgttcag	cccgaaccgt	gcgcccttgc	cggtaactat	cgtcttgagt	ccaacccggt	5100
aagacacgac	ttatcgccac	tggcagcagc	cactggtaac	aggattagca	gagcgaggta	5160
tgtaggcggg	gctacagagt	tcttgaagtg	gtggcctaac	tacggctaca	ctagaagaac	5220
agtatttggg	atctgcgctc	tgctgaagcc	agttaccttc	ggaaaaagag	ttggtagctc	5280
ttgatccggc	aaacaaacca	ccgctggtag	cgggtggttt	tttgtttgca	agcagcagat	5340
tacgcgcaga	aaaaaaggat	ctcaagaaga	tcctttgatc	ttttctacgg	ggtctgacgc	5400
tcagtggaac	gaaaactcac	gttaagggat	tttggtcatt	agattatcaa	aaaggatcct	5460
cacctagatc	cttttaaatt	aaaaatgaag	ttttaaatca	atctaaagta	tatatgagta	5520
aacttgggtc	gacagttacc	aatgcttaat	cagtgaggca	cctatctcag	cgatctgtct	5580
atttcgttca	tccatagtgt	cctgactccc	cgtcgtgtag	ataactacga	tacgggaggg	5640
cttaccatct	ggccccagtg	ctgcaatgat	accgcgagac	ccacgctcac	cggctccaga	5700
tttatcagca	ataaaccagc	cagccggaag	ggccgagcgc	agaagtggtc	ctgcaacttt	5760
atccgcctcc	atccagtcta	ttaattgttg	ccgggaagct	agagtaagta	gttcgccagt	5820
taatagtttg	cgcaacgttg	ttgccattgc	tacaggcatc	gtggtgtcac	gctcgtcggt	5880
tggtatgggt	tcattcagct	ccggttccca	acgatcaagg	cgagttacat	gatcccccat	5940
gttgtgcaaa	aaagcggtta	gctccttcgg	tcctccgatc	gttgtcagaa	gtaagttggc	6000
cgcagtgtta	tcactcatgg	ttatggcagc	actgcataat	tctcttactg	tcattgccatc	6060
cgtaagatgc	ttttctgtga	ctggtgagta	ctcaaccaag	tcattctgag	aatagtgtat	6120
gcggcgaccg	agttgctctt	gcccggcgct	aatacgggat	aataccgctc	cacatagcag	6180
aactttaaaa	gtgctcatca	ttggaaaacg	ttcttcgggg	cgaaaactct	caaggatcct	6240
accgctgttg	agatccagtt	cgatgtaacc	cactcgtgca	cccaactgat	cttcagcatc	6300
ttttactttc	accagcgttt	ctgggtgagc	aaaaacagga	aggcaaaatg	ccgcaaaaaa	6360
gggaataaag	gcgacacgga	aatgttgaat	actcatactc	ttcctttttc	aatattattg	6420
aagcatttat	caggggttatt	gtctcatgag	cggatacata	tttgaatgta	tttagaaaaa	6480
taaacaaata	ggggttccgc	gcacatttcc	ccgaaaagtg	ccacctgaac	gaagcatctg	6540
tgcttcattt	tgtagaacaa	aatgcaacg	cgagagcgct	aatttttcaa	acaaagaatc	6600
tgagctgcat	ttttacagaa	cagaaatgca	acgcgaaagc	gctatttttac	caacgaagaa	6660
tctgtgcttc	atttttgtaa	aacaaaaatg	caacgcgaga	gcgctaattt	ttcaaacaaa	6720
gaatctgagc	tgcattttta	cagaacagaa	atgcaacgcg	agagcgctat	tttaccaaca	6780

aagaatctat acttcttttt tgttctacaa aaatgcatcc cgagagcgct atttttctaa	6840
caaagcatct tagattactt tttttctcct ttgtgcgctc tataatgcag tctcttgata	6900
actttttgca ctgtaggtcc gttaagggtta gaagaaggct actttggtgt ctattttctc	6960
ttccataaaa aaagcctgac tccacttccc gcgtttactg attactagcg aagctgcggg	7020
tgcatttttt caagataaag gcatccccga ttatatctta taccgatgtg gattgcgcat	7080
actttgtgaa cagaaagtga tagcgttgat gattcttcat tggtcagaaa attatgaacg	7140
gtttcttcta ttttgtctct atatactacg tataggaaat gtttacattt tcgtattgtt	7200
ttcgattcac tctatgaata gttcttacta caattttttt gtctaaagag taatactaga	7260
gataaacata aaaaatgtag aggtcgagtt tagatgcaag ttcaaggagc gaaagggtga	7320
tgggtaggtt atatagggat atagcacaga gatatatagc aaagagatac ttttgagcaa	7380
tgtttgtgga agcgggtattc gcaatatatt agtagctcgt tacagtccgg tgcgtttttg	7440
gttttttgaa agtgcgtctt cagagcgctt ttggttttca aaagcgctct gaagttccta	7500
tactttctag agaataggaa cttcggaata ggaacttcaa agcgtttccg aaaacgagcg	7560
cttccgaaaa tgcaacgcga gctgcgcaca tacagctcac tgttcacgtc gcacctatat	7620
ctgcgtgttg cctgtatata tatatacatg agaagaacgg catagtgcgt gtttatgctt	7680
aaatgcgtac ttatatgcgt ctatttatgt aggatgaaag gtagtctagt acctcctgtg	7740
atattatccc attccatgcg gggatatgta tgcttccttc agcactaccc tttagctgtt	7800
ctatatgctg ccactcctca attggattag tctcatcctt caatgctatc atttcctttg	7860
atattggatc atactaagaa accattatta tcatgacatt aacctataaa aataggcgta	7920
tcacgaggcc ctttcgtc	7938