

(19) **DANMARK**



Patent- og
Varemærkestyrelsen

(12)

Oversættelse af europæisk patentskrift

(10) **DK/EP 2751261 T3**

-
- (51) Int.Cl.: **C 12 N 9/04 (2006.01)** **C 12 P 7/42 (2006.01)** **C 12 P 7/52 (2006.01)**
- (45) Oversættelsen bekendtgjort den: **2018-12-10**
- (80) Dato for Den Europæiske Patentmyndigheds bekendtgørelse om meddelelse af patentet: **2018-08-29**
- (86) Europæisk ansøgning nr.: **12772176.9**
- (86) Europæisk indleveringsdag: **2012-09-25**
- (87) Den europæiske ansøgnings publiceringsdag: **2014-07-09**
- (86) International ansøgning nr.: **US2012057134**
- (87) Internationalt publikationsnr.: **WO2013049073**
- (30) Prioritet: **2011-09-30 US 201161541363 P**
- (84) Designerede stater: **AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR**
- (73) Patenthaver: **Novozymes, Inc., 1445 Drew Avenue, Davis, CA 95618, USA**
Novozymes A/S, Krogshøjvej 36, 2880 Bagsværd, Danmark
- (72) Opfinder: **TASSONE, Monica, 3348 Paumanok Way, Sacramento, California 95835, USA**
DE MARIA, Leonardo, Schlegels Alle 6, 1807 Frederiksberg C, Danmark
- (74) Fuldmægtig i Danmark: **Novozymes A/S, Krogshøjvej 36, 2880 Bagsværd, Danmark**
- (54) Benævnelse: **Dehydrogenasevarianter og polynukleotider kodende herfor**
- (56) Fremdragne publikationer:
WO-A1-2008/027742
WO-A2-2005/118719
US-A1- 2011 201 073
DATABASE UniProt [Online] EMBL; 29 April 2008 (2008-04-29), "SubName: Full=NADP-dependent l-serine/l-allo-threonine dehydrogenase ydfg; EC=1.1.1.-"; XP002688443, retrieved from EBI accession no. UNIPROT:B1ELX0 Database accession no. B1ELX0
DATABASE UniProt [Online] 21 September 2011 (2011-09-21), "SubName: Full=Putative oxidoreductase;", XP002688444, retrieved from EBI accession no. UNIPROT:F8VES4 Database accession no. F8VES4
DATABASE UniProt [Online] 5 April 2011 (2011-04-05), "SubName: Full=NADP(+)-dependent dehydrogenase;", XP002688445, retrieved from EBI accession no. UNIPROT:E7R2D7 Database accession no. E7R2D7
DATABASE UniProt [Online] 16 August 2004 (2004-08-16), "SubName: Full=KLLA0B08371p;", XP002688446, retrieved from EBI accession no. UNIPROT:Q6CVY7 Database accession no. Q6CVY7
XINGLIN JIANG ET AL: "Biosynthetic pathways for 3-hydroxypropionic acid production", APPLIED MICROBIOLOGY AND BIOTECHNOLOGY, SPRINGER, BERLIN, DE, vol. 82, no. 6, 17 February 2009 (2009-02-17), pages 995-1003, XP019705542, ISSN: 1432-0614
FUJISAWA H ET AL: "Characterization of short-chain dehydrogenase/reductase homologues of Escherichia coli (YdfG) and Saccharomyces cerevisiae (YMR226C)", BIOCHIMICA ET BIOPHYSICA ACTA (BBA) -

Fortsættes ...

PROTEINS & PROTEOMICS, ELSEVIER, NETHERLANDS, vol. 1645, no. 1, 31 January 2003 (2003-01-31), pages 89-94, XP004401463, ISSN: 1570-9639, DOI: 10.1016/S1570-9639(02)00533-2
DATABASE Geneseq [Online] 29 September 2011 (2011-09-29), "Kluyveromyces lactis 3-keto acid reductase homolog protein, SEQ ID 7.", retrieved from EBI accession no. GSP:AZM00759 Database accession no. AZM00759

DESCRIPTION

Background

[0001] 3-hydroxypropionic acid (3-HP) is a three carbon carboxylic acid identified by the U.S. Department of Energy as one of the top 12 high-potential building block chemicals that can be made by fermentation. Alternative names for 3-HP, which is an isomer of lactic (2-hydroxypropionic) acid, include ethylene lactic acid and 3-hydroxypropionate. 3-HP is an attractive renewable platform chemical, with 100% theoretical yield from glucose, multiple functional groups that allow it to participate in a variety of chemical reactions, and low toxicity. 3-HP can be used as a substrate to form several commodity chemicals, such as 1,3-propanediol, malonic acid, acrylamide, and acrylic acid. Acrylic acid is a large-volume chemical (>7 billion lbs/year) used to make acrylate esters and superabsorbent polymers, and is currently derived from catalytic oxidation of propylene. Fermentative production of 3-HP would provide a sustainable alternative to petrochemicals as the feedstock for these commercially-significant chemicals, thus reducing energy consumption, dependence on foreign oil supplies, and the production of greenhouse gases.

[0002] 3-hydroxypropionate dehydrogenase (3-HPDH) is an enzyme that converts malonate semialdehyde to 3-HP (Figure 1). Certain 3-HPDH enzymes utilize the cofactor NADP(H) (EC 1.1.1.298). However, it may be desirable with some engineered metabolic pathways for 3-HPDH to utilize the cofactor NAD(H) rather than NADP(H) (e.g., to improve redox balance). Accordingly, there is a need in the art to develop dehydrogenase variants that have increased specificity for the cofactor NAD(H) compared to NADP(H). Described herein are dehydrogenase variants that meet this need.

[0003] Database entries UniProt B1ELX0 (2008-04-29), F8VES4 (2011-09-21), E7R2D7 (2011-04-05), and Q6CVY7 (2004-08-16), disclose protein sequences translated from genomic entries and predicts possible functions. There is no information that these sequences are expressed in host cells. Database entry GSP: AZM00759 (2011-09-29), discloses the sequence of an expressed protein having less than 80% identity to the sequences recited in the claims. WO 2008/027742 A1 discloses the use of *E. coli* YfdG (present Seq ID No 2) in the production of 3-hydroxypropionic acid. The difference with the present mutated enzymes is that they have an increased specificity for NADH when compared to NADPH.

Summary

[0004] In a first aspect, the present invention relates to a host cell comprising an active 3-hydroxypropionate pathway and a heterologous polynucleotide encoding a 3-hydroxypropionate dehydrogenase variant having 3-hydroxypropionate dehydrogenase activity, wherein the variant comprises or consists of SEQ ID NO: 10, 12, 20, 21, 22, 23, 24, 25, 26, 27, 29, 30, 32, or 81.

[0005] In a second aspect, the present invention relates to a method of producing 3-hydroxypropionate, comprising:

1. a) cultivating the host cell of the previous aspect under conditions conducive for production of 3-hydroxypropionate; and
2. b) recovering the 3-hydroxypropionate.

Brief Description of the Figures

[0006]

Figure 1 shows a pathway for generating 3-HP.

Figure 2 shows an alignment of native dehydrogenase sequences for *E. coli* ydfG, *I. orientalis* YMR226c, and *S. cerevisiae* YMR226c (SEQ ID NOs: 2, 4, and 6, respectively). Residues involved in cofactor binding are underlined. Residues involved in catalysis are boldfaced.

Figure 3 shows a partial sequence alignment for the N-terminal region of variant dehydrogenases mut1-mut25 (SEQ ID NOs: 7, 8, 9, 10, 11, 12, 13, 14, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, and 33, respectively) compared to the native *E. coli* dehydrogenase (SEQ ID NO: 2); and for the N-terminal region of variant dehydrogenases mut26 and mut27 (SEQ ID NOs: 80 and 81, respectively) compared to the native *I. orientalis* dehydrogenase (SEQ ID NO: 4).

Figure 4 shows plasmid map for pTrc99A.

Figure 5 shows plasmid map for pMcTs108.

Figure 6 shows plasmid map for pMcTs116.

Figure 7 shows plasmid map for p1045168.

Figure 8 shows plasmid map for pMcTs77.

Figure 9 shows plasmid map for p11AAT5WP.

Figure 10 shows plasmid map for pMcTs78.

Figure 11 shows plasmid map for pMcTs115.

Figure 12 shows plasmid map for p11AA2GJP.

Figure 13 shows plasmid map for pMcTs102.

Definitions

[0007] 3-hydroxypropionate dehydrogenase: The term "3-hydroxypropionate dehydrogenase" (3-HPDH) means an enzyme that catalyzes the interconversion of malonate semialdehyde to 3-hydroxypropionate (3-HP) in the presence of a NAD(H) or NADP(H) cofactor. Enzymes having 3-HP dehydrogenase activity are classified as EC 1.1.1.59 if they utilize an NAD(H) cofactor, and as EC 1.1.1.298 if they utilize an NADP(H) cofactor. Enzymes classified as EC 1.1.1.298 are alternatively referred to as malonate semialdehyde reductases. One skilled in the art will recognize that 3-hydroxypropionate dehydrogenases may have specificity for more than one substrate. For example, the *E. coli* 3-hydroxypropionate dehydrogenase of SEQ ID NO: 2 may catalyze both the interconversion of serine to 2-aminomalonate semialdehyde (i.e. a "serine dehydrogenase") and the interconversion of 3-HP to malonate semialdehyde (i.e., a 3-HPDH).

[0008] 3-hydroxypropionate dehydrogenase activity can be determined according to malonate semialdehyde reductase assay described in the Examples. In one aspect, the variants of the present invention have at least 20%, e.g., at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at

least 95%, or at least 100% of the 3-hydroxypropionate dehydrogenase of SEQ ID NO: 2, 4, or 6.

[0009] Active 3-hydroxypropionate pathway: As used herein, a host cell having an "active 3-HP pathway" produces active enzymes necessary to catalyze each reaction in a metabolic pathway from a fermentable sugar to 3-HP, and therefore is capable of producing 3-HP in measurable yields when cultivated under fermentation conditions in the presence of at least one fermentable sugar. A host cell having an active 3-HP pathway comprises one or more 3-HP pathway genes. A "3-HP pathway gene" as used herein refers to a gene that encodes an enzyme involved in an active 3-HP pathway. One example of an active 3-HP pathway and corresponding enzymes involved in the active 3-HP pathway is shown in Figure 1.

[0010] The active enzymes necessary to catalyze each reaction in active 3-HP pathway may result from activities of endogenous gene expression, activities of heterologous gene expression, or from a combination of activities of endogenous and heterologous gene expression.

[0011] Coding sequence: The term "coding sequence" means a polynucleotide sequence, which specifies the amino acid sequence of a polypeptide. The boundaries of the coding sequence are generally determined by an open reading frame, which usually begins with the ATG start codon or alternative start codons such as GTG and TTG and ends with a stop codon such as TAA, TAG, and TGA. The coding sequence may be a sequence of genomic DNA, cDNA, a synthetic polynucleotide, and/or a recombinant polynucleotide.

[0012] Control sequence: The term "control sequence" means a nucleic acid sequence necessary for polypeptide expression. Control sequences may be native or foreign to the polynucleotide encoding the polypeptide, and native or foreign to each other. Such control sequences include, but are not limited to, a leader sequence, polyadenylation sequence, propeptide sequence, promoter sequence, signal peptide sequence, and transcription terminator sequence. The control sequences may be provided with linkers for the purpose of introducing specific restriction sites facilitating ligation of the control sequences with the coding region of the polynucleotide encoding a polypeptide.

[0013] Expression: The term "expression" includes any step involved in the production of the polypeptide including, but not limited to, transcription, post-transcriptional modification, translation, post-translational modification, and secretion. Expression can be measured-for example, to detect increased expression-by techniques known in the art, such as measuring levels of mRNA and/or translated polypeptide.

[0014] Expression vector: The term "expression vector" means a linear or circular DNA molecule that comprises a polynucleotide encoding a polypeptide and is operably linked to control sequences, wherein the control sequences provide for expression of the polynucleotide encoding the polypeptide. At a minimum, the expression vector comprises a promoter sequence, and transcriptional and translational stop signal sequences.

[0015] Fermentable medium: The term "fermentable medium" refers to a medium comprising one or more (e.g., two, several) sugars, such as glucose, fructose, sucrose, cellobiose, xylose, xylulose, arabinose, mannose, galactose, and/or soluble oligosaccharides, wherein the medium is capable, in part, of being converted (fermented) into 3-HP by a host cell having an active 3-HP pathway. In some instances, the fermentation medium is derived from a natural source, such as sugar cane, starch, or cellulose, and may be the result of pretreating the source by enzymatic hydrolysis (saccharification).

[0016] Fragment: The term "fragment" means a polypeptide having one or more (e.g., two, several) amino acids deleted from the amino and/or carboxyl terminus of a referenced polypeptide sequence. In one

aspect, the fragment has 3-HPDH activity. In another aspect, the number of amino acid residues in the fragment is at least 75%, e.g., at least 80%, 85%, 90%, or 95% of any 3-HPDH herein, e.g., at least 75%, e.g., at least 80%, 85%, 90%, or 95% of the number of amino acid residues in SEQ ID NOs: 2, 4, or 6.

[0017] Heterologous polynucleotide: The term "heterologous polynucleotide" is defined herein as a polynucleotide that is not native to the host cell; a native polynucleotide in which one or more (e.g., two, several) structural modifications have been made to the coding region; a native polynucleotide whose expression is quantitatively altered as a result of manipulation of the DNA by recombinant DNA techniques, e.g., a different (foreign) promoter linked to the polynucleotide; or a native polynucleotide whose expression is quantitatively altered by the introduction of one or more extra copies of the polynucleotide into the host cell.

[0018] Host cell: The term "host cell" means any cell type that is susceptible to transformation, transfection, transduction, and the like with a nucleic acid construct or expression vector comprising a polynucleotide described herein (e.g., a polynucleotide encoding a 3-HPDH). The term "host cell" encompasses any progeny of a parent cell that is not identical to the parent cell due to mutations that occur during replication.

[0019] Increased specificity: The term "increased specificity for NAD(H) compared to NADP(H)" means the referenced polypeptide has greater 3-HPDH activity in the presence of NAD(H) compared to NADP(H) in otherwise identical conditions. In some aspects, the referenced variant has more than 2-fold, e.g., more than 5-fold, 10-fold, 20-fold, 50-fold, 100-fold, 200-fold, 500-fold, or 1000-fold specificity for NAD(H) compared to NADP(H).

[0020] Isolated: The term "isolated" means a substance in a form or environment which does not occur in nature. Non-limiting examples of isolated substances include (1) any non-naturally occurring substance, (2) any substance including, but not limited to, any enzyme, variant, nucleic acid, protein, peptide or cofactor, that is at least partially removed from one or more or all of the naturally occurring constituents with which it is associated in nature; (3) any substance modified by the hand of man relative to that substance found in nature; or (4) any substance modified by increasing the amount of the substance relative to other components with which it is naturally associated (e.g., multiple copies of a gene encoding the substance; use of a stronger promoter than the promoter naturally associated with the gene encoding the substance). An isolated substance may be present in a fermentation broth sample.

[0021] Mutant: The term "mutant" means a polynucleotide encoding a variant.

[0022] Nucleic acid construct: The term "nucleic acid construct" means a polynucleotide that comprises one or more (e.g., two, several) control sequences. The polynucleotide may be single-stranded or double-stranded, and may be isolated from a naturally occurring gene, modified to contain segments of nucleic acids in a manner that would not otherwise exist in nature, or synthetic.

[0023] Operably linked: The term "operably linked" means a configuration in which a control sequence is placed at an appropriate position relative to the coding sequence of a polynucleotide such that the control sequence directs the expression of the coding sequence.

[0024] Sequence identity: The relatedness between two amino acid sequences or between two nucleotide sequences is described by the parameter "sequence identity".

[0025] For purposes of the present invention, the sequence identity between two amino acid sequences is determined using the Needleman-Wunsch algorithm (Needleman and Wunsch, 1970, J. Mol. Biol. 48: 443-

453) as implemented in the Needle program of the EMBOSS package (EMBOSS: The European Molecular Biology Open Software Suite, Rice et al., 2000, Trends Genet. 16: 276-277), preferably version 5.0.0 or later. The parameters used are gap open penalty of 10, gap extension penalty of 0.5, and the EBLOSUM62 (EMBOSS version of BLOSUM62) substitution matrix.

[0026] For purposes of the present invention, the sequence identity between two deoxyribonucleotide sequences is determined using the Needleman-Wunsch algorithm (Needleman and Wunsch, 1970, *supra*) as implemented in the Needle program of the EMBOSS package (EMBOSS: The European Molecular Biology Open Software Suite, Rice et al., 2000, *supra*), preferably version 5.0.0 or later. The parameters used are gap open penalty of 10, gap extension penalty of 0.5, and the EDNAFULL (EMBOSS version of NCBI NUC4.4) substitution matrix.

[0027] Subsequence: The term "subsequence" means a polynucleotide having one or more (e.g., two, several) nucleotides deleted from the 5' and/or 3' end of the referenced nucleotide sequence. In one aspect, the subsequence encodes a fragment having 3-HPDH activity. In another aspect, the number of nucleotide residues in the subsequence is at least 75%, e.g., at least 80%, 85%, 90%, or 95% of the number of nucleotide residues in any sequence encoding a 3-HPDH described herein, e.g., at least 75%, e.g., at least 80%, 85%, 90%, or 95% of the number of nucleotide residues in SEQ ID NOs: 1, 3, or 5.

[0028] Variant: The term "variant" means a 3-HPDH comprising an alteration, *i.e.*, a substitution, insertion, and/or deletion, at one or more (e.g., two, several) positions relative to a parent 3-HPDH. A substitution means replacement of the amino acid occupying a position with a different amino acid; a deletion means removal of the amino acid occupying a position; and an insertion means adding an amino acid adjacent to and immediately following the amino acid occupying a position.

[0029] Wild-type: The term "wild-type" 3-HPDH or "native" 3-HPDH means a 3-HPDH expressed by a naturally occurring microorganism, such as a bacterium, yeast, or filamentous fungus found in nature.

Conventions for Designation of Variants

[0030] For purposes described herein, SEQ ID NO: 2 is used to determine amino acid numbering in other 3-HPDH enzymes. The amino acid sequence of another 3-HPDH is aligned with SEQ ID NO: 2, and based on the alignment, the amino acid position number corresponding to any amino acid residue in the polypeptide disclosed in SEQ ID NO: 2 is determined using the Needleman-Wunsch algorithm (Needleman and Wunsch, 1970, J. Mol. Biol. 48: 443-453) as implemented in the Needle program of the EMBOSS package (EMBOSS: The European Molecular Biology Open Software Suite, Rice et al., 2000, Trends Genet. 16: 276-277), preferably version 5.0.0 or later. The parameters used are gap open penalty of 10, gap extension penalty of 0.5, and the EBLOSUM62 (EMBOSS version of BLOSUM62) substitution matrix.

[0031] Identification of the corresponding amino acid residue in another 3-HPDH can be determined by an alignment of multiple polypeptide sequences using several computer programs including, but not limited to, MUSCLE (multiple sequence comparison by log-expectation; version 3.5 or later; Edgar, 2004, Nucleic Acids Research 32: 1792-1797), MAFFT (version 6.857 or later; Katoh and Kuma, 2002, Nucleic Acids Research 30: 3059-3066; Katoh et al., 2005, Nucleic Acids Research 33: 511-518; Katoh and Toh, 2007, Bioinformatics 23: 372-374; Katoh et al., 2009, Methods in Molecular Biology 537: 39-64; Katoh and Toh, 2010, Bioinformatics 26: 1899-1900), and EMBOSS EMMA employing ClustalW (1.83 or later; Thompson et al., 1994, Nucleic Acids Research 22: 4673-4680), using their respective default parameters.

[0032] When the other enzyme sequence has diverged from the SEQ ID NO: 2 such that traditional

sequence-based comparison fails to detect their relationship (Lindahl and Elofsson, 2000, J. Mol. Biol. 295: 613-615), other pairwise sequence comparison algorithms can be used. Greater sensitivity in sequence-based searching can be attained using search programs that utilize probabilistic representations of polypeptide families (profiles) to search databases. For example, the PSI-BLAST program generates profiles through an iterative database search process and is capable of detecting remote homologs (Atschul et al., 1997, Nucleic Acids Res. 25: 3389-3402). Even greater sensitivity can be achieved if the family or superfamily for the polypeptide has one or more representatives in the protein structure databases. Programs such as GenTHREADER (Jones, 1999, J. Mol. Biol. 287: 797-815; McGuffin and Jones, 2003, Bioinformatics 19: 874-881) utilize information from a variety of sources (PSI-BLAST, secondary structure prediction, structural alignment profiles, and solvation potentials) as input to a neural network that predicts the structural fold for a query sequence. Similarly, the method of Gough et al., 2000, J. Mol. Biol. 313: 903-919, can be used to align a sequence of unknown structure with the superfamily models present in the SCOP database. These alignments can in turn be used to generate homology models for the polypeptide, and such models can be assessed for accuracy using a variety of tools developed for that purpose.

[0033] For proteins of known structure, several tools and resources are available for retrieving and generating structural alignments. For example the SCOP superfamilies of proteins have been structurally aligned, and those alignments are accessible and downloadable. Two or more protein structures can be aligned using a variety of algorithms such as the distance alignment matrix (Holm and Sander, 1998, Proteins 33: 88-96) or combinatorial extension (Shindyalov and Bourne, 1998, Protein Engineering 11: 739-747), and implementation of these algorithms can additionally be utilized to query structure databases with a structure of interest in order to discover possible structural homologs (e.g., Holm and Park, 2000, Bioinformatics 16: 566-567).

[0034] In describing the variants described herein, the nomenclature described below is adapted for ease of reference. The accepted IUPAC single letter or three letter amino acid abbreviation is employed.

[0035] Substitutions. For an amino acid substitution, the following nomenclature is used: Original amino acid, position, substituted amino acid. Accordingly, the substitution of threonine at position 226 with alanine is designated as "Thr226Ala" or "T226A". Alternative substitutions at the same position are separated by a slant. For example, the substitution of threonine at position 226 with alanine or valine is designated as "Thr226Ala/Val" or "T226A/V", representing a T226A or T226V substitution. Multiple mutations are separated by addition marks ("+"), e.g., "Gly205Arg + Ser411Phe/Tyr" or "G205R + S411F/Y", representing substitutions at positions 205 and 411 of glycine (G) with arginine (R) and serine (S) with phenylalanine (F) or tyrosine (Y), respectively.

[0036] Deletions. For an amino acid deletion, the following nomenclature is used: Original amino acid, position, *. Accordingly, the deletion of glycine at position 195 is designated as "Gly195*" or "G195*". Multiple deletions are separated by addition marks ("+"), e.g., "Gly195* + Ser411*" or "G195* + S411*".

[0037] Insertions. For an amino acid insertion, the following nomenclature is used: Original amino acid, position, original amino acid, inserted amino acid. Accordingly the insertion of lysine after glycine at position 195 is designated "Gly195GlyLys" or "G195GK". An insertion of multiple amino acids is designated [Original amino acid, position, original amino acid, inserted amino acid #1, inserted amino acid #2; etc.]. For example, the insertion of lysine and alanine after glycine at position 195 is indicated as "Gly195GlyLysAla" or "G195GKA".

[0038] In such cases the inserted amino acid residue(s) are numbered by the addition of lower case letters to the position number of the amino acid residue preceding the inserted amino acid residue(s). In the above example, the sequence would thus be:

Parent:	Variant:
195	195 195a 195b
G	G - K - A

[0039] Reference to "about" a value or parameter herein includes aspects that are directed to that value or parameter *per se*. For example, description referring to "about X" includes the aspect "X". When used in combination with measured values, "about" includes a range that encompasses at least the uncertainty associated with the method of measuring the particular value, and can include a range of plus or minus two standard deviations around the stated value.

[0040] As used herein and in the appended claims, the singular forms "a," "or," and "the" include plural referents unless the context clearly dictates otherwise. It is understood that the aspects described herein include "consisting" and/or "consisting essentially of" aspects.

[0041] Unless defined otherwise or clearly indicated by context, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art.

Detailed Description

[0042] Described herein, *inter alia*, are host cells comprising an active 3-hydroxypropionate pathway and a heterologous polynucleotide encoding a 3-hydroxypropionate dehydrogenase variant having 3-hydroxypropionate dehydrogenase activity, wherein the variant comprises or consists of SEQ ID NO: 10, 12, 20, 21, 22, 23, 24, 25, 26, 27, 29, 30, 32, or 81.

[S (BvW1)]

Host Cells

[0043] The present invention relates to recombinant host cells comprising an active 3-hydroxypropionate pathway and a heterologous polynucleotide encoding a 3-hydroxypropionate dehydrogenase variant having 3-hydroxypropionate dehydrogenase activity, wherein the variant comprises or consists of SEQ ID NO: 10, 12, 20, 21, 22, 23, 24, 25, 26, 27, 29, 30, 32, or 81. A construct or vector comprising a heterologous polynucleotide is introduced into a host cell so that the construct or vector is maintained as a chromosomal integrant or as a self-replicating extrachromosomal vector as described earlier. The term "host cell" encompasses any progeny of a parent cell that is not identical to the parent cell due to mutations that occur during replication.

[0044] In some aspects, the 3-hydroxypropionate dehydrogenase variant has increased specificity for NAD(H) compared to NADP(H) (e.g., greater than 2-fold, 5-fold, 10-fold, 20-fold, 50-fold, 100-fold, 200-fold, 500-fold, or 1000-fold specificity for NAD(H) compared to NADP(H)).

[0045] In some aspects, a host cell may be selected for the recombinant production and recovery of the 3-hydroxypropionate dehydrogenase variant. In other aspects, the host cell comprises an active 3-hydroxypropionate pathway and is chosen to express the variant as a 3-hydroxypropionate pathway gene in the recombinant production of 3-hydroxypropionate by the cell (e.g., as described in WO2012/074818). Such cells can produce 3-hydroxypropionate from a fermentable sugar or a malonyl semialdehyde precursor.

[0046] The prokaryotic host cell may be any Gram-positive or Gram-negative bacterium. Gram-positive bacteria include, but are not limited to, *Bacillus*, *Clostridium*, *Enterococcus*, *Geobacillus*, *Lactobacillus*, *Lactococcus*, *Oceanobacillus*, *Staphylococcus*, *Streptococcus*, and *Streptomyces*. Gram-negative bacteria include, but are not limited to, *Campylobacter*, *E. coli*, *Flavobacterium*, *Fusobacterium*, *Helicobacter*, *Ilyobacter*, *Neisseria*, *Pseudomonas*, *Salmonella*, and *Ureaplasma*.

[0047] The bacterial host cell may be any *Bacillus* cell including, but not limited to, *Bacillus alkalophilus*, *Bacillus amyloliquefaciens*, *Bacillus brevis*, *Bacillus circulans*, *Bacillus clausii*, *Bacillus coagulans*, *Bacillus firmus*, *Bacillus lautus*, *Bacillus lentus*, *Bacillus licheniformis*, *Bacillus megaterium*, *Bacillus pumilus*, *Bacillus stearothermophilus*, *Bacillus subtilis*, and *Bacillus thuringiensis* cells.

[0048] The bacterial host cell may also be any *Streptococcus* cell including, but not limited to, *Streptococcus equisimilis*, *Streptococcus pyogenes*, *Streptococcus uberis*, and *Streptococcus equi* subsp. *Zooepidemicus* cells.

[0049] The bacterial host cell may also be any *Streptomyces* cell, including, but not limited to, *Streptomyces achromogenes*, *Streptomyces avermitilis*, *Streptomyces coelicolor*, *Streptomyces griseus*, and *Streptomyces lividans* cells.

[0050] The introduction of DNA into a *Bacillus* cell may be effected by protoplast transformation (see, e.g., Chang and Cohen, 1979, Mol. Gen. Genet. 168: 111-115), competent cell transformation (see, e.g., Young and Spizizen, 1961, J. Bacteriol. 81: 823-829, or Dubnau and Davidoff-Abelson, 1971, J. Mol. Biol. 56: 209-221), electroporation (see, e.g., Shigekawa and Dower, 1988, Biotechniques 6: 742-751), or conjugation (see, e.g., Koehler and Thorne, 1987, J. Bacteriol. 169: 5271-5278). The introduction of DNA into an *E. coli* cell may be effected by protoplast transformation (see, e.g., Hanahan, 1983, J. Mol. Biol. 166: 557-580) or electroporation (see, e.g., Dower et al., 1988, Nucleic Acids Res. 16: 6127-6145). The introduction of DNA into a *Streptomyces* cell may be effected by protoplast transformation, electroporation (see, e.g., Gong et al., 2004, Folia Microbiol. (Praha) 49: 399-405), conjugation (see, e.g., Mazodier et al., 1989, J. Bacteriol. 171: 3583-3585), or transduction (see, e.g., Burke et al., 2001, Proc. Natl. Acad. Sci. USA 98: 6289-6294). The introduction of DNA into a *Pseudomonas* cell may be effected by electroporation (see, e.g., Choi et al., 2006, J. Microbiol. Methods 64: 391-397), or conjugation (see, e.g., Pinedo and Smets, 2005, Appl. Environ. Microbiol. 71: 51-57). The introduction of DNA into a *Streptococcus* cell may be effected by natural competence (see, e.g., Perry and Kuramitsu, 1981, Infect. Immun. 32: 1295-1297), protoplast transformation (see, e.g., Catt and Jollick, 1991, Microbios 68: 189-207), electroporation (see, e.g., Buckley et al., 1999, Appl. Environ. Microbiol. 65: 3800-3804) or conjugation (see, e.g., Clewell, 1981, Microbiol. Rev. 45: 409-436). However, any method known in the art for introducing DNA into a host cell can be used.

[0051] The host cell may also be a eukaryote, such as a mammalian, insect, plant, or fungal cell.

[0052] The host cell may be a fungal cell. "Fungi" as used herein includes the phyla Ascomycota, Basidiomycota, Chytridiomycota, and Zygomycota as well as the Oomycota and all mitosporic fungi (as defined by Hawksworth et al., In, Ainsworth and Bisby's Dictionary of The Fungi, 8th edition, 1995, CAB International, University Press, Cambridge, UK).

[0053] The fungal host cell may be a yeast cell. "Yeast" as used herein includes ascosporogenous yeast (Endomycetales), basidiosporogenous yeast, and yeast belonging to the Fungi Imperfecti (Blastomycetes). Since the classification of yeast may change in the future, for the purposes of this invention, yeast shall be defined as described in Biology and Activities of Yeast (Skinner, Passmore, and Davenport, editors, Soc. App. Bacteriol. Symposium Series No. 9, 1980).

[0054] The yeast host cell may be a *Candida*, *Hansenula*, *Kluyveromyces*, *Pichia*, *Saccharomyces*, *Schizosaccharomyces*, or *Yarrowia* cell such as a *Kluyveromyces lactis*, *Saccharomyces carlsbergensis*, *Saccharomyces cerevisiae*, *Saccharomyces diastaticus*, *Saccharomyces douglasii*, *Saccharomyces kluyveri*, *Saccharomyces norbensis*, *Saccharomyces oviformis*, or *Yarrowia lipolytica* cell.

[0055] In some aspects, the host cell is selected from *Issatchenkia*, *Candida*, *Kluyveromyces*, *Pichia*, *Schizosaccharomyces*, *Torulaspora*, *Zygosaccharomyces*, and *Saccharomyces*. In some aspects, the host cell is a *I. orientalis*, *C. lambica*, or *S. bulderi* host cell.

[0056] The fungal host cell may be a filamentous fungal cell. "Filamentous fungi" include all filamentous forms of the subdivision Eumycota and Oomycota (as defined by Hawksworth et al., 1995, *supra*). The filamentous fungi are generally characterized by a mycelial wall composed of chitin, cellulose, glucan, chitosan, mannan, and other complex polysaccharides. Vegetative growth is by hyphal elongation and carbon catabolism is obligately aerobic. In contrast, vegetative growth by yeasts such as *Saccharomyces cerevisiae* is by budding of a unicellular thallus and carbon catabolism may be fermentative.

[0057] The filamentous fungal host cell may be an *Acremonium*, *Aspergillus*, *Aureobasidium*, *Bjerkandera*, *Ceriporiopsis*, *Chrysosporium*, *Coprinus*, *Coriolus*, *Cryptococcus*, *Filibasidium*, *Fusarium*, *Humicola*, *Magnaporthe*, *Mucor*, *Myceliophthora*, *Neocallimastix*, *Neurospora*, *Paecilomyces*, *Penicillium*, *Phanerochaete*, *Phlebia*, *Piromyces*, *Pleurotus*, *Schizophyllum*, *Talaromyces*, *Thermoascus*, *Thielavia*, *Tolypocladium*, *Trametes*, or *Trichoderma* cell.

[0058] For example, the filamentous fungal host cell may be an *Aspergillus awamori*, *Aspergillus foetidus*, *Aspergillus fumigatus*, *Aspergillus japonicus*, *Aspergillus nidulans*, *Aspergillus niger*, *Aspergillus oryzae*, *Bjerkandera adusta*, *Ceriporiopsis aneirina*, *Ceriporiopsis caregiea*, *Ceriporiopsis gilvescens*, *Ceriporiopsis pannocinta*, *Ceriporiopsis rivulosa*, *Ceriporiopsis subrufa*, *Ceriporiopsis subvermispora*, *Chrysosporium inops*, *Chrysosporium keratinophilum*, *Chrysosporium lucknowense*, *Chrysosporium merdarium*, *Chrysosporium pannicola*, *Chrysosporium queenslandicum*, *Chrysosporium tropicum*, *Chrysosporium zonatum*, *Coprinus cinereus*, *Coriolus hirsutus*, *Fusarium bactridioides*, *Fusarium cerealis*, *Fusarium crookwellense*, *Fusarium culmorum*, *Fusarium graminearum*, *Fusarium graminum*, *Fusarium heterosporum*, *Fusarium negundi*, *Fusarium oxysporum*, *Fusarium reticulatum*, *Fusarium roseum*, *Fusarium sambucinum*, *Fusarium sarcochroum*, *Fusarium sporotrichioides*, *Fusarium sulphureum*, *Fusarium torulosum*, *Fusarium trichothecioides*, *Fusarium venenatum*, *Humicola insolens*, *Humicola lanuginosa*, *Mucor miehei*, *Myceliophthora thermophila*, *Neurospora crassa*, *Penicillium purpurogenum*, *Phanerochaete chrysosporium*, *Phlebia radiata*, *Pleurotus eryngii*, *Thielavia terrestris*, *Trametes villosa*, *Trametes versicolor*, *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei*, or *Trichoderma viride* cell.

[0059] Fungal cells may be transformed by a process involving protoplast formation, transformation of the protoplasts, and regeneration of the cell wall in a manner known *per se*. Suitable procedures for transformation of *Aspergillus* and *Trichoderma* host cells are described in EP 238023, Yelton et al., 1984, Proc. Natl. Acad. Sci. USA 81: 1470-1474, and Christensen et al., 1988, Bio/Technology 6: 1419-1422. Suitable methods for transforming *Fusarium* species are described by Malardier et al., 1989, Gene 78: 147-156, and WO 96/00787. Yeast may be transformed using the procedures described by Becker and Guarente, In Abelson, J.N. and Simon, M.I., editors, Guide to Yeast Genetics and Molecular Biology, Methods in Enzymology, Volume 194, pp 182-187, Academic Press, Inc., New York; Ito et al., 1983, J. Bacteriol. 153: 163; and Hinnen et al., 1978, Proc. Natl. Acad. Sci. USA 75: 1920.

[0060] In some aspects, the host cell comprises an active 3-hydroxypropionate pathway which includes a

heterologous polynucleotide that encodes a 3-hydroxypropionate dehydrogenase variant described herein and is capable of producing 3-hydroxypropionate from a fermentable sugar (e.g., glucose) or pyruvate. Active 3-hydroxypropionate pathways such as the pathway shown in Figure 1 are known in the art (see, for example, WO2012/074818). The host cell may comprise PEP carboxylase activity or pyruvate carboxylase activity; aspartate aminotransferase activity; aspartate decarboxylase activity; and beta-alanine/alpha-ketoglutarate aminotransferase (BAAT) activity. (see, for example, WO02/42418 and WO2008/027742). Such enzyme activities may result from endogenous gene expression, expression of heterologous polynucleotides encoding the enzymes in the metabolic pathway, or from a combination of endogenous gene expression supplemented with expression of one or more (e.g., two, several) heterologous polynucleotides. In some embodiments, the host cell comprises a heterologous polynucleotide that encodes a PEP carboxylase, a heterologous polynucleotide that encodes a pyruvate carboxylase, a heterologous polynucleotide that encodes a aspartate aminotransferase, a heterologous polynucleotide that encodes a aspartate decarboxylase, and/or a heterologous polynucleotide that encodes a BAAT.

[0061] In some aspects, the host cell is a 3-hydroxypropionate-resistant host cell. A "3-hydroxypropionate-resistant host cell" as used herein refers to a host cell that exhibits an average glycolytic rate of at least 2.5 g/L/hr in media containing 75 g/L or greater 3-HP at a pH of less than 4.0. Such rates and conditions represent an economic process for producing 3-HP. In certain of these embodiments, the host cells may exhibit 3-hydroxypropionate resistance in their native form. In other embodiments, the host cells may have undergone mutation and/or selection before, during, or after introduction of genetic modifications related to an active 3-hydroxypropionate fermentation pathway, such that the mutated and/or selected cells possess a higher degree of resistance to 3-hydroxypropionate than wild-type cells of the same species. In certain embodiments, mutation and/or selection may be carried out on cells that exhibit 3-hydroxypropionate resistance in their native form. Cells that have undergone mutation and/or selection may be tested for sugar consumption and other characteristics in the presence of varying levels of 3-hydroxypropionate in order to determine their potential as industrial hosts for 3-hydroxypropionate production. In addition to 3-hydroxypropionate resistance, the host cells provided herein may have undergone mutation and/or selection for resistance to one or more additional organic acids or to other fermentation products, byproducts, or media components.

[0062] Selection for resistance to 3-hydroxypropionate or to other compounds may be accomplished using methods well known in the art. For example, selection may be carried out using a chemostat. A chemostat is a device that allows for a continuous culture of microorganisms (e.g., yeast) wherein the specific growth rate and cell number can be controlled independently. A continuous culture is essentially a flow system of constant volume to which medium is added continuously and from which continuous removal of any overflow can occur. Once such a system is in equilibrium, cell number and nutrient status remain constant, and the system is in a steady state. A chemostat allows control of both the population density and the specific growth rate of a culture through dilution rate and alteration of the concentration of a limiting nutrient, such as a carbon or nitrogen source. By altering the conditions as a culture is grown (e.g., decreasing the concentration of a secondary carbon source necessary to the growth of the inoculum strain, among others), microorganisms in the population that are capable of growing faster at the altered conditions will be selected and will outgrow microorganisms that do not function as well under the new conditions. Typically such selection requires the progressive increase or decrease of at least one culture component over the course of growth of the chemostat culture. The operation of chemostats and their use in the directed evolution of microorganisms is well known in the art (see, e.g., Novick *Proc Natl Acad Sci USA* 36:708-719 (1950), Harder *J Appl Bacteriol* 43:1-24 (1977)).

[0063] In some aspects, the host cell secretes (and/or is capable of secreting) an increased level of 3-hydroxypropionate compared to the host cell without the heterologous polynucleotide that encodes the 3-hydroxypropionate dehydrogenase variant described herein when cultivated under the same conditions. In

some embodiments, the host cell secretes and/or is capable of secreting an increased level of 3-hydroxypropionate of at least 5%, e.g., at least 10%, at least 15%, at least 20%, at least 25%, at least 50%, at least 100%, at least 150%, at least 200%, at least 300%, or at 500% compared to the host cell without the heterologous polynucleotide that encodes the 3-hydroxypropionate dehydrogenase variant described herein, when cultivated under the same conditions. Examples of suitable cultivation conditions are described below and will be readily apparent to one of skill in the art based on the teachings herein. In some embodiments, the host cell produces (and/or is capable of producing) 3-hydroxypropionate at a yield of at least 10%, e.g., at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, or at least 90%, of theoretical. In some embodiments, the host cell has a 3-hydroxypropionate volumetric productivity greater than about 0.1 g/L per hour, e.g., greater than about 0.2 g/L per hour, 0.5 g/L per hour, 0.6 g/L per hour, 0.7 g/L per hour, 0.8 g/L per hour, 0.9 g/L per hour, 1.0 g/L per hour, 1.1 g/L per hour, 1.2 g/L per hour, 1.3 g/L per hour, 1.5 g/L per hour, 1.75 g/L per hour, 2.0 g/L per hour, 2.25 g/L per hour, 2.5 g/L per hour, or 3.0 g/L per hour; or between about 0.1 g/L per hour and about 2.0 g/L per hour, e.g., between about 0.3 g/L per hour and about 1.7 g/L per hour, about 0.5 g/L per hour and about 1.5 g/L per hour, about 0.7 g/L per hour and about 1.3 g/L per hour, about 0.8 g/L per hour and about 1.2 g/L per hour, or about 0.9 g/L per hour and about 1.1 g/L per hour.

[0064] The host cells may be cultivated in a nutrient medium suitable for production of the 3-hydroxypropionate dehydrogenase variants described herein using methods well known in the art. For example, the cell may be cultivated by shake flask cultivation, and small-scale or large-scale fermentation (including continuous, batch, fed-batch, or solid state fermentations) in laboratory or industrial fermentors performed in a suitable medium and under conditions allowing the desired polypeptide to be expressed and/or isolated. The cultivation takes place in a suitable nutrient medium comprising carbon and nitrogen sources and inorganic salts, as described herein, using procedures known in the art. Suitable media are available from commercial suppliers, may be prepared according to published compositions (e.g., in catalogues of the American Type Culture Collection), or may be prepared from commercially available ingredients.

[0065] As described supra, enzyme activities of the enzymes described herein can be detected using methods known in the art. These detection methods may include use of specific antibodies, formation of an enzyme product, or disappearance of an enzyme substrate. See, for example, Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Third Ed., Cold Spring Harbor Laboratory, New York (2001); Ausubel et al., *Current Protocols in Molecular Biology*, John Wiley and Sons, Baltimore, MD (1999); and Hanai et al., *Appl. Environ. Microbiol.* 73:7814-7818 (2007)).

Preparation of Variants

[0066] Also described are methods for preparing a 3-hydroxypropionate dehydrogenase variant having 3-hydroxypropionate dehydrogenase activity, wherein the variant comprises or consists of SEQ ID NO: 10, 12, 20, 21, 22, 23, 24, 25, 26, 27, 29, 30, 32, or 81.

[0067] The variants can be prepared using any mutagenesis procedure known in the art, such as site-directed mutagenesis, synthetic gene construction, semi-synthetic gene construction, random mutagenesis, shuffling, etc.

[0068] Site-directed mutagenesis is a technique in which one or more (e.g., several) mutations are introduced at one or more defined sites in a polynucleotide encoding the parent.

[0069] Site-directed mutagenesis can be accomplished *in vitro* by PCR involving the use of oligonucleotide primers containing the desired mutation. Site-directed mutagenesis can also be performed *in vitro* by cassette mutagenesis involving the cleavage by a restriction enzyme at a site in the plasmid comprising a polynucleotide encoding the parent and subsequent ligation of an oligonucleotide containing the mutation in the polynucleotide. Usually the restriction enzyme that digests the plasmid and the oligonucleotide is the same, permitting sticky ends of the plasmid and the insert to ligate to one another. See, e.g., Scherer and Davis, 1979, Proc. Natl. Acad. Sci. USA 76: 4949-4955; and Barton et al., 1990, Nucleic Acids Res. 18: 7349-4966.

[0070] Site-directed mutagenesis can also be accomplished *in vivo* by methods known in the art. See, e.g., U.S. Patent Application Publication No. 2004/0171154; Storici et al., 2001, Nature Biotechnol. 19: 773-776; Kren et al., 1998, Nat. Med. 4: 285-290; and Calissano and Macino, 1996, Fungal Genet. Newslett. 43: 15-16.

[0071] Any site-directed mutagenesis procedure can be used to prepare the variants described herein. For example, there are many commercial kits available that can be used.

[0072] Synthetic gene construction entails *in vitro* synthesis of a designed polynucleotide molecule to encode a polypeptide of interest. Gene synthesis can be performed utilizing a number of techniques, such as the multiplex microchip-based technology described by Tian et al. (2004, Nature 432: 1050-1054) and similar technologies wherein oligonucleotides are synthesized and assembled upon photo-programmable microfluidic chips.

[0073] Single or multiple amino acid substitutions, deletions, and/or insertions can be made and tested using known methods of mutagenesis, recombination, and/or shuffling, followed by a relevant screening procedure, such as those disclosed by Reidhaar-Olson and Sauer, 1988, Science 241: 53-57; Bowie and Sauer, 1989, Proc. Natl. Acad. Sci. USA 86: 2152-2156; WO 95/17413; or WO 95/22625. Other methods that can be used include error-prone PCR, phage display (e.g., Lowman et al., 1991, Biochemistry 30: 10832-10837; U.S. Patent No. 5,223,409; WO 92/06204) and region-directed mutagenesis (Derbyshire et al., 1986, Gene 46: 145; Ner et al., 1988, DNA 7: 127).

[0074] Mutagenesis/shuffling methods can be combined with high-throughput, automated screening methods to detect activity of cloned, mutagenized polypeptides expressed by host cells (Ness et al., 1999, Nature Biotechnology 17: 893-896). Mutagenized DNA molecules that encode active polypeptides can be recovered from the host cells and rapidly sequenced using standard methods in the art. These methods allow the rapid determination of the importance of individual amino acid residues in a polypeptide.

[0075] Semi-synthetic gene construction is accomplished by combining aspects of synthetic gene construction, and/or site-directed mutagenesis, and/or random mutagenesis, and/or shuffling. Semi-synthetic construction is typified by a process utilizing polynucleotide fragments that are synthesized, in combination with PCR techniques. Defined regions of genes may thus be synthesized *de novo*, while other regions may be amplified using site-specific mutagenic primers, while yet other regions may be subjected to error-prone PCR or non-error prone PCR amplification. Polynucleotide subsequences may then be shuffled.

Polynucleotides, Nucleic Acid Constructs, and Expression Vectors

[0076] In another aspect are heterologous polynucleotides encoding 3-hydroxypropionate dehydrogenase variants having 3-hydroxypropionate dehydrogenase activity as described herein, as well as nucleic acid constructs and expression vectors comprising the heterologous polynucleotides.

[0077] The nucleic acid constructs comprise a heterologous polynucleotide encoding a variant described herein operably linked to one or more control sequences that direct the expression of the coding sequence in a suitable host cell under conditions compatible with the control sequences.

[0078] The heterologous polynucleotide may be manipulated in a variety of ways to provide for expression of a variant. Manipulation of the heterologous polynucleotide prior to its insertion into a vector may be desirable or necessary depending on the expression vector. The techniques for modifying polynucleotides utilizing recombinant DNA methods are well known in the art.

[0079] The control sequence may be a promoter, a polynucleotide which is recognized by a host cell for expression of the heterologous polynucleotide. The promoter contains transcriptional control sequences that mediate the expression of the variant. The promoter may be any polynucleotide that shows transcriptional activity in the host cell including mutant, truncated, and hybrid promoters, and may be obtained from genes encoding extracellular or intracellular polypeptides either homologous or heterologous to the host cell.

[0080] Examples of suitable promoters for directing transcription of the nucleic acid constructs of the present invention in a bacterial host cell are the promoters obtained from the *Bacillus amyloliquefaciens* alpha-amylase gene (*amyQ*), *Bacillus licheniformis* alpha-amylase gene (*amyL*), *Bacillus licheniformis* penicillinase gene (*penP*), *Bacillus stearothermophilus* maltogenic amylase gene (*amyM*), *Bacillus subtilis* levansucrase gene (*sacB*), *Bacillus subtilis* *xylA* and *xylB* genes, *Bacillus thuringiensis* *cryIIIA* gene (Agaisse and Lereclus, 1994, Molecular Microbiology 13: 97-107), *E. coli* *lac* operon, *E. coli* *trc* promoter (Egon et al., 1988, Gene 69: 301-315), *Streptomyces coelicolor* agarase gene (*dagA*), and prokaryotic beta-lactamase gene (Villa-Kamaroff et al., 1978, Proc. Natl. Acad. Sci. USA 75: 3727-3731), as well as the *tac* promoter (DeBoer et al., 1983, Proc. Natl. Acad. Sci. USA 80: 21-25). Further promoters are described in "Useful proteins from recombinant bacteria" in Gilbert et al., 1980, Scientific American 242: 74-94; and in Sambrook et al., 1989, *supra*. Examples of tandem promoters are disclosed in WO 99/43835.

[0081] Examples of suitable promoters for directing transcription of the nucleic acid constructs of the present invention in a filamentous fungal host cell are promoters obtained from the genes for *Aspergillus nidulans* acetamidase, *Aspergillus niger* neutral alpha-amylase, *Aspergillus niger* acid stable alpha-amylase, *Aspergillus niger* or *Aspergillus awamori* glucoamylase (*glaA*), *Aspergillus oryzae* TAKA amylase, *Aspergillus oryzae* alkaline protease, *Aspergillus oryzae* triose phosphate isomerase, *Fusarium oxysporum* trypsin-like protease (WO 96/00787), *Fusarium venenatum* amyloglucosidase (WO 00/56900), *Fusarium venenatum* Daria (WO 00/56900), *Fusarium venenatum* Quinn (WO 00/56900), *Rhizomucor miehei* lipase, *Rhizomucor miehei* aspartic proteinase, *Trichoderma reesei* beta-glucosidase, *Trichoderma reesei* cellobiohydrolase I, *Trichoderma reesei* cellobiohydrolase II, *Trichoderma reesei* endoglucanase I, *Trichoderma reesei* endoglucanase II, *Trichoderma reesei* endoglucanase III, *Trichoderma reesei* endoglucanase IV, *Trichoderma reesei* endoglucanase V, *Trichoderma reesei* xylanase I, *Trichoderma reesei* xylanase II, *Trichoderma reesei* beta-xylosidase, as well as the NA2-tpi promoter (a modified promoter from an *Aspergillus* neutral alpha-amylase gene in which the untranslated leader has been replaced by an untranslated leader from an *Aspergillus* triose phosphate isomerase gene; non-limiting examples include modified promoters from an *Aspergillus niger* neutral alpha-amylase gene in which the untranslated leader has been replaced by an untranslated leader from an *Aspergillus nidulans* or *Aspergillus oryzae* triose phosphate isomerase gene); and mutant, truncated, and hybrid promoters thereof.

[0082] In a yeast host, useful promoters are obtained from the genes for *Saccharomyces cerevisiae* enolase (ENO-1), *Saccharomyces cerevisiae* galactokinase (GAL1), *Saccharomyces cerevisiae* alcohol 3-hydroxypropionate dehydrogenase/glyceraldehyde-3-phosphate 3-hydroxypropionate dehydrogenase

(ADH1, ADH2/GAP), *Saccharomyces cerevisiae* triose phosphate isomerase (TPI), *Saccharomyces cerevisiae* metallothionein (CUP1), and *Saccharomyces cerevisiae* 3-phosphoglycerate kinase. Other useful promoters for yeast host cells are described by Romanos et al., 1992, *Yeast* 8: 423-488.

[0083] The control sequence may also be a transcription terminator, which is recognized by a host cell to terminate transcription. The terminator sequence is operably linked to the 3'-terminus of the heterologous polynucleotide encoding the variant. Any terminator that is functional in the host cell may be used.

[0084] Preferred terminators for bacterial host cells are obtained from the genes for *Bacillus clausii* alkaline protease (*aprH*), *Bacillus licheniformis* alpha-amylase (*amyL*), and *Escherichia coli* ribosomal RNA (*rrnB*).

[0085] Preferred terminators for filamentous fungal host cells are obtained from the genes for *Aspergillus nidulans* anthranilate synthase, *Aspergillus niger* glucoamylase, *Aspergillus niger* alpha-glucosidase, *Aspergillus oryzae* TAKA amylase, and *Fusarium oxysporum* trypsin-like protease.

[0086] Preferred terminators for yeast host cells are obtained from the genes for *Saccharomyces cerevisiae* enolase, *Saccharomyces cerevisiae* cytochrome C (CYC1), and *Saccharomyces cerevisiae* glyceraldehyde-3-phosphate 3-hydroxypropionate dehydrogenase. Other useful terminators for yeast host cells are described by Romanos et al., 1992, *supra*.

[0087] The control sequence may also be an mRNA stabilizer region downstream of a promoter and upstream of the coding sequence of a gene which increases expression of the gene.

[0088] Examples of suitable mRNA stabilizer regions are obtained from a *Bacillus thuringiensis cryIIIA* gene (WO 94/25612) and a *Bacillus subtilis* SP82 gene (Hue et al., 1995, *Journal of Bacteriology* 177: 3465-3471).

[0089] The control sequence may also be a leader, a nontranslated region of an mRNA that is important for translation by the host cell. The leader sequence is operably linked to the 5'-terminus of the heterologous polynucleotide encoding the variant. Any leader that is functional in the host cell may be used.

[0090] Preferred leaders for filamentous fungal host cells are obtained from the genes for *Aspergillus oryzae* TAKA amylase and *Aspergillus nidulans* triose phosphate isomerase.

[0091] Suitable leaders for yeast host cells are obtained from the genes for *Saccharomyces cerevisiae* enolase (ENO-1), *Saccharomyces cerevisiae* 3-phosphoglycerate kinase, *Saccharomyces cerevisiae* alpha-factor, and *Saccharomyces cerevisiae* alcohol 3-hydroxypropionate dehydrogenase/glyceraldehyde-3-phosphate 3-hydroxypropionate dehydrogenase (ADH2/GAP).

[0092] The control sequence may also be a polyadenylation sequence, a sequence operably linked to the 3'-terminus of the variant-encoding sequence and, when transcribed, is recognized by the host cell as a signal to add polyadenosine residues to transcribed mRNA. Any polyadenylation sequence that is functional in the host cell may be used.

[0093] Preferred polyadenylation sequences for filamentous fungal host cells are obtained from the genes for *Aspergillus nidulans* anthranilate synthase, *Aspergillus niger* glucoamylase, *Aspergillus niger* alpha-glucosidase, *Aspergillus oryzae* TAKA amylase, and *Fusarium oxysporum* trypsin-like protease.

[0094] Useful polyadenylation sequences for yeast host cells are described by Guo and Sherman, 1995, *Mol. Cellular Biol.* 15: 5983-5990.

[0095] The control sequence may also be a signal peptide coding region that encodes a signal peptide linked to the N-terminus of a variant and directs the variant into the cell's secretory pathway. The 5'-end of the coding sequence of the heterologous polynucleotide may inherently contain a signal peptide coding sequence naturally linked in translation reading frame with the segment of the coding sequence that encodes the variant. Alternatively, the 5'-end of the coding sequence may contain a signal peptide coding sequence that is foreign to the coding sequence. A foreign signal peptide coding sequence may be required where the coding sequence does not naturally contain a signal peptide coding sequence. Alternatively, a foreign signal peptide coding sequence may simply replace the natural signal peptide coding sequence in order to enhance secretion of the variant. However, any signal peptide coding sequence that directs the expressed variant into the secretory pathway of a host cell may be used.

[0096] Effective signal peptide coding sequences for bacterial host cells are the signal peptide coding sequences obtained from the genes for *Bacillus* NCIB 11837 maltogenic amylase, *Bacillus licheniformis* subtilisin, *Bacillus licheniformis* beta-lactamase, *Bacillus stearothermophilus* alpha-amylase, *Bacillus stearothermophilus* neutral proteases (*nprT*, *nprS*, *nprM*), and *Bacillus subtilis* *prsA*. Further signal peptides are described by Simonen and Palva, 1993, Microbiological Reviews 57: 109-137.

[0097] Effective signal peptide coding sequences for filamentous fungal host cells are the signal peptide coding sequences obtained from the genes for *Aspergillus niger* neutral amylase, *Aspergillus niger* glucoamylase, *Aspergillus oryzae* TAKA amylase, *Humicola insolens* cellulase, *Humicola insolens* endoglucanase V, *Humicola lanuginosa* lipase, and *Rhizomucor miehei* aspartic proteinase.

[0098] Useful signal peptides for yeast host cells are obtained from the genes for *Saccharomyces cerevisiae* alpha-factor and *Saccharomyces cerevisiae* invertase. Other useful signal peptide coding sequences are described by Romanos et al., 1992, *supra*.

[0099] The control sequence may also be a propeptide coding sequence that encodes a propeptide positioned at the N-terminus of a variant. The resultant polypeptide is known as a proenzyme or propolypeptide (or a zymogen in some cases). A propolypeptide is generally inactive and can be converted to an active polypeptide by catalytic or autocatalytic cleavage of the propeptide from the propolypeptide. The propeptide coding sequence may be obtained from the genes for *Bacillus subtilis* alkaline protease (*aprE*), *Bacillus subtilis* neutral protease (*nprT*), *Myceliophthora thermophila* laccase (WO 95/33836), *Rhizomucor miehei* aspartic proteinase, and *Saccharomyces cerevisiae* alpha-factor.

[0100] Where both signal peptide and propeptide sequences are present, the propeptide sequence is positioned next to the N-terminus of the variant and the signal peptide sequence is positioned next to the N-terminus of the propeptide sequence.

[0101] It may also be desirable to add regulatory sequences that regulate expression of the variant relative to the growth of the host cell. Examples of regulatory systems are those that cause expression of the gene to be turned on or off in response to a chemical or physical stimulus, including the presence of a regulatory compound. Regulatory systems in prokaryotic systems include the *lac*, *tac*, and *trp* operator systems. In yeast, the ADH2 system or GAL1 system may be used. In filamentous fungi, the *Aspergillus niger* glucoamylase promoter, *Aspergillus oryzae* TAKA alpha-amylase promoter, and *Aspergillus oryzae* glucoamylase promoter may be used. Other examples of regulatory sequences are those that allow for gene amplification. In eukaryotic systems, these regulatory sequences include the dihydrofolate reductase gene that is amplified in the presence of methotrexate, and the metallothionein genes that are amplified with heavy metals. In these cases, the heterologous polynucleotide encoding the variant would be operably linked with the regulatory sequence.

[0102] Recombinant expression vectors comprise a heterologous polynucleotide encoding a variant described herein, a promoter, and transcriptional and translational stop signals. The various nucleotide and control sequences may be joined together to produce a recombinant expression vector that may include one or more convenient restriction sites to allow for insertion or substitution of the heterologous polynucleotide encoding the variant at such sites. Alternatively, the heterologous polynucleotide may be expressed by inserting the heterologous polynucleotide or a nucleic acid construct comprising the heterologous polynucleotide into an appropriate vector for expression. In creating the expression vector, the coding sequence is located in the vector so that the coding sequence is operably linked with the appropriate control sequences for expression.

[0103] The recombinant expression vector may be any vector (e.g., a plasmid or virus) that can be conveniently subjected to recombinant DNA procedures and can bring about expression of the heterologous polynucleotide. The choice of the vector will typically depend on the compatibility of the vector with the host cell into which the vector is to be introduced. The vector may be a linear or closed circular plasmid.

[0104] The vector may be an autonomously replicating vector, *i.e.*, a vector that exists as an extrachromosomal entity, the replication of which is independent of chromosomal replication, e.g., a plasmid, an extrachromosomal element, a minichromosome, or an artificial chromosome. The vector may contain any means for assuring self-replication. Alternatively, the vector may be one that, when introduced into the host cell, is integrated into the genome and replicated together with the chromosome(s) into which it has been integrated. Furthermore, a single vector or plasmid or two or more vectors or plasmids that together contain the total DNA to be introduced into the genome of the host cell, or a transposon, may be used.

[0105] The vector preferably contains one or more selectable markers that permit easy selection of transformed, transfected, transduced, or the like cells. A selectable marker is a gene the product of which provides for biocide or viral resistance, resistance to heavy metals, prototrophy to auxotrophs, and the like.

[0106] Examples of bacterial selectable markers are *Bacillus licheniformis* or *Bacillus subtilis* *dal* genes, or markers that confer antibiotic resistance such as ampicillin, chloramphenicol, kanamycin, neomycin, spectinomycin or tetracycline resistance. Suitable markers for yeast host cells include, but are not limited to, ADE2, HIS3, LEU2, LYS2, MET3, TRP1, and URA3. Selectable markers for use in a filamentous fungal host cell include, but are not limited to, *amdS* (acetamidase), *argB* (ornithine carbamoyltransferase), *bar* (phosphinothricin acetyltransferase), *hph* (hygromycin phosphotransferase), *niaD* (nitrate reductase), *pyrG* (orotidine-5'-phosphate decarboxylase), *sC* (sulfate adenylyltransferase), and *trpC* (anthranilate synthase), as well as equivalents thereof. Preferred for use in an *Aspergillus* cell are *Aspergillus nidulans* or *Aspergillus oryzae* *amdS* and *pyrG* genes and a *Streptomyces hygroscopicus* *bar* gene.

[0107] The vector preferably contains an element(s) that permits integration of the vector into the host cell's genome or autonomous replication of the vector in the cell independent of the genome.

[0108] For integration into the host cell genome, the vector may rely on the heterologous polynucleotide's sequence encoding the variant or any other element of the vector for integration into the genome by homologous or non-homologous recombination. Alternatively, the vector may contain additional polynucleotides for directing integration by homologous recombination into the genome of the host cell at a precise location(s) in the chromosome(s). To increase the likelihood of integration at a precise location, the integrational elements should contain a sufficient number of nucleic acids, such as 100 to 10,000 base pairs, 400 to 10,000 base pairs, and 800 to 10,000 base pairs, which have a high degree of sequence identity to the corresponding target sequence to enhance the probability of homologous recombination. The integrational elements may be any sequence that is homologous with the target sequence in the genome of

the host cell. Furthermore, the integrational elements may be non-encoding or encoding polynucleotides. On the other hand, the vector may be integrated into the genome of the host cell by non-homologous recombination.

[0109] For autonomous replication, the vector may further comprise an origin of replication enabling the vector to replicate autonomously in the host cell in question. The origin of replication may be any plasmid replicator mediating autonomous replication that functions in a cell. The term "origin of replication" or "plasmid replicator" means a polynucleotide that enables a plasmid or vector to replicate *in vivo*.

[0110] Examples of bacterial origins of replication are the origins of replication of plasmids pBR322, pUC19, pACYC177, and pACYC184 permitting replication in *E. coli*, and pUB110, pE194, pTA1060, and pAM β 1 permitting replication in *Bacillus*.

[0111] Examples of origins of replication for use in a yeast host cell are the 2 micron origin of replication, ARS1, ARS4, the combination of ARS1 and CEN3, and the combination of ARS4 and CEN6.

[0112] Examples of origins of replication useful in a filamentous fungal cell are AMA1 and ANS1 (Gems et al., 1991, Gene 98: 61-67; Cullen et al., 1987, Nucleic Acids Res. 15: 9163-9175; WO 00/24883). Isolation of the AMA1 gene and construction of plasmids or vectors comprising the gene can be accomplished according to the methods disclosed in WO 00/24883.

[0113] More than one copy of a heterologous polynucleotide of the present invention may be inserted into a host cell to increase production of a variant. An increase in the copy number of the heterologous polynucleotide can be obtained by integrating at least one additional copy of the sequence into the host cell genome or by including an amplifiable selectable marker gene with the heterologous polynucleotide where cells containing amplified copies of the selectable marker gene, and thereby additional copies of the heterologous polynucleotide, can be selected for by cultivating the cells in the presence of the appropriate selectable agent.

[0114] The procedures used to ligate the elements described above to construct the recombinant expression vectors of the present invention are well known to one skilled in the art (see, e.g., Sambrook et al., 1989, *supra*).

Methods of producing 3-HP

[0115] The host cells described herein may be used for the production of 3-hydroxypropionate. In one aspect is a method of producing 3-hydroxypropionate from a fermentable sugar (e.g., glucose) or pyruvate, comprising: (a) cultivating any one of the host cells described herein (e.g., a host cell that comprises an active 3-hydroxypropionate pathway and a heterologous polynucleotide that encodes a 3-hydroxypropionate dehydrogenase variant described herein) in a medium under suitable conditions to produce the 3-hydroxypropionate; and (b) recovering the 3-hydroxypropionate. In some embodiments of the method, the host cell further comprises PEP carboxylase activity or pyruvate carboxylase activity; aspartate aminotransferase activity; aspartate decarboxylase activity; and beta-alanine/alpha-ketoglutarate aminotransferase (BAAT) activity. In some embodiments, the host cell comprises a heterologous polynucleotide that encodes a PEP carboxylase, a heterologous polynucleotide that encodes a pyruvate carboxylase, a heterologous polynucleotide that encodes an aspartate aminotransferase, a heterologous polynucleotide that encodes an aspartate decarboxylase, and/or a heterologous polynucleotide that encodes a BAAT. In some embodiments of the methods, the host cells are 3-hydroxypropionate resistant host cells, as described *supra*.

[0116] Methods for the production of 3-hydroxypropionate may be performed in a fermentable medium comprising any one or more (e.g., two, several) sugars, such as glucose, fructose, sucrose, cellobiose, xylose, xylulose, arabinose, mannose, galactose, and/or soluble oligosaccharides. In some instances, the fermentation medium is derived from a natural source, such as sugar cane, starch, or cellulose, and may be the result of pretreating the source by enzymatic hydrolysis (saccharification).

[0117] In addition to the appropriate carbon sources from one or more (e.g., two, several) sugar(s), the fermentable medium may contain other nutrients or stimulators known to those skilled in the art, such as macronutrients (e.g., nitrogen sources) and micronutrients (e.g., vitamins, mineral salts, and metallic cofactors). In some aspects, the carbon source can be preferentially supplied with at least one nitrogen source, such as yeast extract, N₂, peptone (e.g., Bacto™ Peptone), or soytone (e.g., Bacto™ Soytone). Nonlimiting examples of vitamins include multivitamins, biotin, pantothenate, nicotinic acid, meso-inositol, thiamine, pyridoxine, para-aminobenzoic acid, folic acid, riboflavin, and Vitamins A, B, C, D, and E. Examples of mineral salts and metallic cofactors include, but are not limited to Na, P, K, Mg, S, Ca, Fe, Zn, Mn, and Cu.

[0118] Suitable conditions used for the methods of 3-hydroxypropionate production may be determined by one skilled in the art in light of the teachings herein. In some aspects of the methods, the host cells are cultivated for about 12 hours to about 216 hours, such as about 24 hours to about 144 hours, or about 36 hours to about 96 hours. The temperature is typically between about 26°C to about 60°C, e.g., about 34°C to about 50°C, and at a pH of about 3.0 to about 8.0, such as about 3.0 to about 7.0, about 3.0 to about 6.0, about 3.0 to about 5.0, about 3.5 to about 4.5, about 4.0 to about 8.0, about 4.0 to about 7.0, about 4.0 to about 6.0, about 4.0 to about 5.0, about 5.0 to about 8.0, about 5.0 to about 7.0, or about 5.0 to about 6.0. In some aspects of the methods, the resulting intracellular pH of the host cell is about 3.0 to about 8.0, such as about 3.0 to about 7.0, about 3.0 to about 6.0, about 3.0 to about 5.0, about 3.5 to about 4.5, about 4.0 to about 8.0, about 4.0 to about 7.0, about 4.0 to about 6.0, about 4.0 to about 5.0, about 5.0 to about 8.0, about 5.0 to about 7.0, or about 5.0 to about 6.0. Cultivation may be performed under anaerobic, microaerobic, or aerobic conditions, as appropriate. In some aspects, the cultivation is performed under anaerobic conditions. Suitable buffering agents are known in the art.

[0119] Cultivation may be performed under anaerobic, substantially anaerobic (microaerobic), or aerobic conditions, as appropriate. Briefly, anaerobic refers to an environment devoid of oxygen, substantially anaerobic (microaerobic) refers to an environment in which the concentration of oxygen is less than air, and aerobic refers to an environment wherein the oxygen concentration is approximately equal to or greater than that of the air. Substantially anaerobic conditions include, for example, a culture, batch fermentation or continuous fermentation such that the dissolved oxygen concentration in the medium remains less than 10% of saturation. Substantially anaerobic conditions also includes growing or resting cells in liquid medium or on solid agar inside a sealed chamber maintained with an atmosphere of less than 1 % oxygen. The percent of oxygen can be maintained by, for example, sparging the culture with an N₂/CO₂ mixture or other suitable non-oxygen gas or gases. In some embodiments, the cultivation is performed under anaerobic conditions or substantially anaerobic conditions.

[0120] The methods of described herein can employ any suitable fermentation operation mode. For example, a batch mode fermentation may be used with a close system where culture media and host microorganism, set at the beginning of fermentation, have no additional input except for the reagents certain reagents, e.g., for pH control, foam control or others required for process sustenance. The process described herein can also be employed in Fed-batch or continuous mode.

[0121] The methods described herein may be practiced in several bioreactor configurations, such as stirred

tank, bubble column, airlift reactor and others known to those skilled in the art.

[0122] The methods may be performed in free cell culture or in immobilized cell culture as appropriate. Any material support for immobilized cell culture may be used, such as alginates, fibrous bed, or argyle materials such as chrysotile, montmorillonite KSF and montmorillonite K-10.

[0123] In one aspect of the methods, the 3-hydroxypropionate is produced at a titer greater than about 10 g/L, e.g., greater than about 25 g/L, 50 g/L, 75 g/L, 100 g/L, 125 g/L, 150 g/L, 160 g/L, 170 g/L, 180 g/L, 190 g/L, 200 g/L, 210 g/L, 225 g/L, 250 g/L, 275 g/L, 300 g/L, 325 g/L, 350 g/L, 400 g/L, or 500g/L; or between about 10 g/L and about 500 g/L, e.g., between about 50 g/L and about 350 g/L, about 100 g/L and about 300 g/L, about 150 g/L and about 250 g/L, about 175 g/L and about 225 g/L, or about 190 g/L and about 210 g/L. In one aspect of the methods, the 3-hydroxypropionate is produced at a titer greater than about 0.01 gram per gram of carbohydrate, e.g., greater than about 0.02, 0.05, 0.75, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, or 1.0 gram per gram of carbohydrate.

[0124] In one aspect of the methods, the amount of produced 3-hydroxypropionate is at least 5%, e.g., at least 10%, at least 15%, at least 20%, at least 25%, at least 30%, at least 50%, or at least 100% greater compared to cultivating the host cell without the heterologous polynucleotide that encodes a 3-hydroxypropionate dehydrogenase variant under the same conditions.

[0125] The recombinant 3-hydroxypropionate can be optionally recovered and purified from the fermentation medium using any procedure known in the art including, but not limited to, chromatography (e.g., size exclusion chromatography, adsorption chromatography, ion exchange chromatography), electrophoretic procedures, differential solubility, distillation, extraction (e.g., liquid-liquid extraction), pervaporation, extractive filtration, membrane filtration, membrane separation, reverse osmosis, ultrafiltration, or crystallization.

[0126] In some aspects of the methods, the recombinant 3-hydroxypropionate before and/or after being optionally purified is substantially pure. With respect to the methods of producing 3-hydroxypropionate, "substantially pure" intends a recovered preparation of 3-hydroxypropionate that contains no more than 15% impurity, wherein impurity intends compounds other than 3-hydroxypropionate. In one variation, a preparation of substantially pure 3-hydroxypropionate is provided wherein the preparation contains no more than 25% impurity, or no more than 20% impurity, or no more than 10% impurity, or no more than 5% impurity, or no more than 3% impurity, or no more than 1% impurity, or no more than 0.5% impurity.

[0127] Suitable assays to test for the production of 3-hydroxypropionate for the methods of production and host cells described herein can be performed using methods known in the art. For example, the final 3-hydroxypropionate (and other organic compounds) can be analyzed by methods such as HPLC (High Performance Liquid Chromatography), GC-MS (Gas Chromatography Mass Spectroscopy) and LC-MS (Liquid Chromatography-Mass Spectroscopy) or other suitable analytical methods using routine procedures well known in the art. The release of 3-hydroxypropionate in the fermentation broth can also be tested with the culture supernatant. Byproducts and residual sugar in the fermentation medium (e.g., glucose) can be quantified by HPLC using, for example, a refractive index detector for glucose and alcohols, and a UV detector for organic acids (Lin et al., *Biotechnol. Bioeng.* 90:775 -779 (2005)), or using other suitable assay and detection methods well known in the art.

[0128] The present invention is further described by the following examples that should not be construed as limiting the scope of the invention.

Examples

[0129] Chemicals used as buffers and substrates were commercial products of at least reagent grade.

Strains

[0130] The *E. coli* strain MG1655 (NN059268) was used as the source of DNA encoding the ydfG 3-HPDH gene. Strains MG1655, SoloPack Gold (Agilent Technologies, Inc., Santa Clara, CA, USA), and SURE cells (Agilent Technologies, Inc.) were used to express the ydfG plasmids.

[0131] The *I. orientalis* strain MBin500 was used as the source of DNA encoding the *I. orientalis* YMR226c 3-HPDH gene, as described in WO2012/074818. *I. orientalis* strain McTs259 (WO2012/074818) was used to express the described 3-HPDH genes for 3-HP production.

Media

[0132] LB medium was composed of 10 g of tryptone, 5 g of yeast extract, 5 g of sodium chloride, and deionized water to 1 liter.

[0133] 2XYT plates were composed of 16 g of tryptone, 10 g of yeast extract, 5 g of NaCl, 15 g of Bacto agar, and deionized water to 1 liter.

Table 0. Primer Sequences

Identifier	SEQ ID	Sequence (5'-3')
000001	15	CGGAATTCATGATCGTTTTAGTAACTGGAGC
000002	16	CGGGATCCTTACTGACGGTGGACATTCAG
614464	34	TCGCCACTGATCTGAACCCGGAAGCGTTGCAGGAGTTAAAAGA
614465	35	TCTTTTAACTCCTGCAACGCTTCCGGGTTTCAGATCAGTGGCGA
614466	36	CCACTGATCTGAACCCGGCCCGGTTGCAGGAGTTAAAAGACGA
614467	37	TCGTCTTTTAACTCCTGCAACCGGGCCGGGTTTCAGATCAGTGG
614468	38	GGCATAAAGTTATCGCCACTGGCCTGAACCCGGCCGCGTTGCA
614469	39	TGCAACGCGGCCGGGTTTCAGGCCAGTGGCGATAACTTTATGCC
614470	40	TCGTTTTAGTAACTGGAGCAACGGCAGGTTTTGGTGAATGCATT
614471	41	AATGCATTACCAAAACCTGCCGTTGCTCCAGTTACTAAAACGA
614472	42	ATAAAGTTATCGCCACTGATCGTAACCCGGCCGCGTTGCAGGA
614473	43	TCCTGCAACGCGGCCGGGTTACGATCAGTGGCGATAACTTTAT
614476	44	AACTCCTGCAACGCGGCCGGGCGCAGATCAGTGGCGATAACTT
614477	45	AAGTTATCGCCACTGATCTGCGCCCGGCCGCGTTGCAGGAGTT
614479	46	TTATCGCCACTGATCTGAACCAGGCCGCGTTGCAGGAGTTAAA
614480	47	TTTAACTCCTGCAACGCGGCCTGGTTCAGATCAGTGGCGATAA
614546	48	TAAAGTTATCGCCACTGATCTGCGCCCGGAAGCGTTGCAGGAGTTAAA GACG
		CGTCTTTTAACTCCTGCAACGCTTCCGGGTTTCAGATCAGTGGCGATAA

Identifier	SEQ ID	Sequence (5'-3')
614547	49	CGTCTTTTAACTCCTGCAACGCTTCGGGCGCAGATCAGTGGCGATAAC TTTA
614548	50	AGTTATCGCCACTGATCTGCGCCAGGCCGCGTTGCAGGAGTTAAAAGAC GAAC
614549	51	GTTCTGCTTTTAACTCCTGCAACGCGGCCTGGCGCAGATCAGTGGCGAT AACT
614550	52	GCATAAAGTTATCGCCACTGATCTGCGCCAGGAAGCGTTGCAGGAGTTA AAAGACGAAC
614551	53	GTTCTGCTTTTAACTCCTGCAACGCTTCCTGGCGCAGATCAGTGGCGAT AACTTTATGC
614552	54	GCATAAAGTTATCGCCACTGATCTGAACCAGGAAGCGTTGCAGGAGTTA AAAG
614553	55	CTTTTAACTCCTGCAACGCTTCCTGGTTCAGATCAGTGGCGATAACTTTA TGC
614697	56	GTCA TCGTAGTCT AG AT AAAATGA TCGTTTTGGTCACCGG
614698	57	GTGCTCCATTAATTAATTATTGTCTGTG
614967	58	GCGGAATTCATGTTTGGTAATATTTCCCAA
614968	59	GATCCCGGGCTATTTATCTAATGATCCTC
614973	60	GTTTTAGTAACTGGAGCAGGCCGAGGTTTTGGTGAATGC
614974	61	GCATTCACCAAAACCTGCGCCTGCTCCAGTTACTAAAAC
614975	62	CATAAAGTTATCGCCACTGATCGTCGCCAGGAACGGTTG
614976	63	CAACCGTTCCTGGCGACGATCAGTGGCGATAACTTTATG
614977	64	TAAAGTTATCGCCACTGATCGTCGCCAGGAAGCGTTGCAG
614978	65	CTGCAACGCTTCCTGGCGACGATCAGTGGCGATAACTTTA
614979	66	ACTGATCTGCGCCAGGAACGGTTGCAGGAGTTAAAAGAC
614980	67	GTCTTTTAACTCCTGCAACCGTTCCTGGCGCAGATCAGT
615004	68	ATCCTAATTACAGGTGCGGGTACTGGTATCGGATACCAT
615005	69	ATGGTATCCGATACCAGTACCCGCACCTGTAATTAGGAT
615006	70	TTGAAGTTGGTTTTGGCTGATTTAAGAAAGGAGAAGCTGGAG
615007	71	CTCCAGCTTCTCCTTTCTTAAATCAGCCAAAACCAACTTCAA
615428	72	TTGCAGGCAAGAACATCCTAATTACAGGTGC
615429	73	GCACCTGTAATTAGGATGTTCTTGCCTGCAA
615485	74	GTAGCTAGCTAAAATGTTTGGTAATATTTCCCA
615486	75	TGCTTAATTAATATTTATCTAATGATCCTC
615890	76	TTGGTCACCGGTGCAGGTGCAGGTTTCGGCGAA

Identifier	SEQ ID	Sequence (5'-3')
615891	77	TTCGCCGAAACCTGCACCTGCACCGGTGACCAA
615892	78	ACAAGGTTATCGCTACCGACTTGAGACAAGAGAGATTGCA
615893	79	TGCAATCTCTCTTGTCTCAAGTCGGTAGCGATAACCTTGT

Example 1: Construction of an expression vector for the *E. coli* ydfG 3-HPDH gene

[0134] The *E. coli* ydfG 3-HPDH coding sequence was amplified by PCR using two synthetic oligonucleotide primers designed to generate an EcoRI restriction site at the 5' end and a BamHI restriction site at the 3' end for integration into pTrc99A (Figure 4; see Amann, E., et al. (1988). "Tightly regulated tac promoter vectors useful for the expression of unfused and fused proteins in Escherichia coli." Gene 69(2): 301-315).

[0135] *E. coli* genomic DNA for PCR was obtained by isolating a single colony of *E. coli* MG1655 from a 2XYT plate and dissolving into 25 µl 1% Triton X-100, 20 mM Tris pH 8.5, 2 mM EDTA (CLS solution), heated at 80°C for 10 minutes and then cooled on ice. Three microliters of this solution used as a template in a PCR reaction further containing 1X Pfx Amplification buffer, fifty picomoles each of primers 000001 and 000002, 0.2mM each of dATP, dGTP, dCTP, and dTTP, and 2.5 units Platinum® Pfx DNA Polymerase (Invitrogen, Carlsbad, CA, USA) in a final volume of 50 µl. The amplification reaction was performed in an EPPENDORF® MASTERCYCLER® 5333 (Eppendorf Scientific, Inc., Westbury, NY, USA) programmed for one cycle at 95°C for 2 minutes; and 25 cycles each at 95°C for 30 seconds, 55°C for 30 seconds, and 72°C for 1 minute. After the 25 cycles, the reaction was incubated at 72°C for 3 minutes and then cooled at 10°C until further processed.

[0136] Five microliters of the PCR reaction mixture was subjected to 1% TBE-agarose gel electrophoresis with ethidium bromide in TBE buffer to identify the desired 765 bp PCR fragment. The PCR fragment from the remaining 45 µl of the PCR reaction mixture was purified using a QIAquick PCR Purification Kit (Qiagen Inc., Valencia, CA, USA). The purified fragment was digested with EcoRI and BamHI (New England Biolabs, Ipswich, MA, USA) and analyzed on a 1% TBE-agarose gel with ethidium bromide.

[0137] The plasmid pTrc99A (*supra*) was digested with EcoRI and BamHI, and the resulting fragments separated by 1% TBE-agarose gel electrophoresis followed by visualization with a DARK READER™ (Clare Chemical Research, Dolores, CO, USA). The desired 4.1 kb fragment was excised from the gel with a disposable razor blade and purified using a QIAquick Gel Extraction Kit (Qiagen, Inc.).

[0138] Cloning of the DNA fragment containing *E. coli* ydfG into pTrc99A was performed using T4 DNA ligase (New England Biolabs). The reaction mixture contained 1X T4 DNA ligase buffer, 1 µl T4 DNA ligase, 1 µl of the pTrc99A EcoRI/BamHI digested DNA fragment above, and 5 µl of the ydfG EcoRI/BamHI digested PCR product above in total volume of 10µl. The reaction mixture was incubated at 16°C overnight and subsequently used to transform SURE competent cells (Agilent Technologies, Inc.) according to manufacturer's instructions. After a recovery period, two 100 µl aliquots from the transformation mixture were plated onto 150 mm 2XYT plates supplemented with 100 µg of ampicillin per ml and incubated overnight at 37°C.

[0139] Recombinant colonies of the transformations were each inoculated into 3 ml of LB medium supplemented with ampicillin (100 µg/ml). Plasmid DNA was prepared from these cultures using a BIOROBOT® 9600 workstation (Qiagen, Inc.) and analyzed by 1% TBE-agarose gel electrophoresis following EcoRI/BamHI digestion. The plasmid DNA from one clone designated pMeJi9 and having the

correct restriction digest pattern was further subjected to sequence analysis to confirm integration of the correct ydfG coding sequence.

Example 2: Construction of *E. coli* ydfG 3-HPDH gene variants

[0140] Synthetic DNA sequences encoding the desired ydfG variants and containing a 5' flanking EcoRI restriction site and 3' NaeI restriction site were provided in plasmid constructs from DNA2.0 (Menlo Park, CA, USA). Each plasmid was digested with EcoRI and NaeI restriction enzymes, and the resulting fragments separated on a 1% TAE agarose gel followed by visualization with the aid of a DARK READER™ (Clare Chemical Research). The desired DNA band containing the ydfG variant encoding sequence was excised from the gel with a disposable razor blade and purified using NucleoSpin Extract II Kit (Machery-Nagel, Düren, Germany).

[0141] Plasmid pMeJi9 (*supra*) was linearized by digestion with EcoRI and NaeI, followed by incubation with Alkaline Phosphatase, Calf Intestinal (CIP) (New England Biolabs) for removal of the 5' phosphate. The resulting mixture was subjected to gel electrophoresis, visualized, and purified as described above to provide the desired 4657bp DNA fragment.

[0142] Cloning of each ydfG variant encoding sequence into linearized pMeJi9 was performed by incubating 1X T4 DNA ligase buffer, 1 µl T4 DNA ligase (New England Biolabs), 2 µl EcoRI/NaeI linearized pMeJi9, and 15 µl of the selected EcoRI/NaeI ydfG variant encoding sequence (in total volume of 20 µl) for 2 hours at room temperature. A 10µl sample of the incubation reaction was used to transform SoloPack® Gold chemically competent cells according to the manufacturer's instructions. After a recovery period, two 100 µl aliquots from the transformation mixture were plated onto 150 mm 2XYT plates supplemented with 100 µg of ampicillin per ml and incubated overnight at 37°C. Putative recombinant clones were selected from the selection plates and plasmid DNA was prepared from each one using a BIOROBOT® 9600 workstation. Clones were analyzed by sequencing. Those plasmids with the correct sequence are shown in Table 1.

Table 1.

Variant Name	SEQ ID	Cloning Plasmid Name	Amino Acid Changes
Mut1	7	pMcTs68	T9G/A10*/G31D/R32L/R33N/Q34P/E35A/R36A
Mut2	8	pMcTs69	T9G/G31 E/R32L/R33N/Q34P/E35A/R36A
Mut3	9	pMcTs70	T9G/A10*/G31 E/R32L/R33N/Q34P/E35A/R36A
Mut4	10	pMcTs71	T9G/G31D/R32L/R33S/Q34A/E35D/R36A
Mut5	11	pMcTs72	T9G/A10*/G31D/R32L/R33S/Q34A/E35D/R36A
Mut6	12	pMcTs73	T9G/G31D/R32L/R33N/Q34P/E35A/R36A
Mut7	13	pMcTs74	T9G/A10*/G31E/R32L/R33S/Q34A/E35D/R36A
Mut8	14	pMcTs75	T9G/G31 E/R32L/R33S/Q34A/E35D/R36A
*represents deletion of the amino acid.			

[0143] Additional variants shown in Table 2 below were constructed using site-directed mutagenesis and the indicated primers in a PCR reaction using QuikChange II XL Site-Directed Mutagenesis Kit (Agilent Technologies, Inc.). The PCR reaction contained 1X Reaction Buffer, 125ng of each primer, 30ng plasmid

DNA template, 1X dNTPs, 1X Quick solution, 2.5U PfuUltra HF DNA polymerase in a final volume of 50µl. The amplification reaction was performed in an EPPENDORF® MASTERCYCLER® 5333 (Eppendorf Scientific, Inc.) programmed for one cycle at 95°C for 3 minutes; and 18 cycles each at 95°C for 50 seconds, 60°C for 50 seconds, and 68°C for 6 minutes. After the 18 cycles, the reaction was incubated at 68°C for 7 minutes and then cooled at 10°C until further processed. To the each PCR reaction 1ul of DpnI was added and incubated at 37°C for 1.5 hours to digest template plasmid DNA.

[0144] From each site directed PCR reaction, 2.5µl of the reaction was transformed into XL10 Gold Super competent cells (Agilent Technologies, Inc.) according to manufacturer's instructions. After a recovery period, two 100 µl aliquots from the transformation mixture were plated onto 150 mm 2XYT plates supplemented with 100 µg of ampicillin per ml and incubated overnight at 37°C.

[0145] Recombinant colonies of the transformations were each inoculated into 3 ml of LB medium supplemented with ampicillin (100 µg/ml). Plasmid DNA was prepared from these cultures using a BIOROBOT® 9600 workstation (Qiagen, Inc.) and subjected to sequence analysis to confirm the site directed mutation in the ydfG coding sequence.

Table 2.

Variant Name	SEQ ID	Cloning Plasmid Name	Amino Acid Changes	Forward Primer	Reverse Primer	Template
Mut9	17	pMcTs79	G31D/R32L/R33N/Q34P/E35A/R36A	614470	614471	pMcTs73
Mut10	18	pMcTs80	T9G/R32L/R33N/Q34P/E35A/R36A	614468	614469	pMcTs73
Mut11	19	pMcTs81	T9G/G31D/R33N/Q34P/E35A/R36A	614472	614473	pMcTs73
Mut12	20	pMcTs82	T9G/G31 D/R32L/Q34P/E35A/R36A	614476	614477	pMcTs73
Mut13	21	pMcTs83	T9G/G31D/R32L/R33N/E35A/R36A	614479	614480	pMcTs73
Mut14	22	pMcTs84	T9G/G31D/R32L/R33N/Q34P/R36A	614464	614465	pMcTs73
Mut15	23	pMcTs85	T9G/G31D/R32L/R33N/Q34P/E35A	614466	614467	pMcTs73
Mut16	24	pMcTs89	T9G/G31D/R32L/R36A	614550	614551	pMcTs73
Mut17	25	pMcTs86	T9G/G31D/R32L/E35A/R36A	614548	614549	pMcTs73
Mut18	26	pMcTs87	T9G/G31D/R32L/Q34P/R36A	614546	614547	pMcTs73
Mut19	27	pMcTs88	T9G/G31D/R32L/R33N/R36A	614552	614553	pMcTs73
Mut20	28	pMcTs98	G31D	614975	614976	pMeJi9
Mut21	29	pMcTs99	T9G/G31D/R36A	614977	614978	pMcTs89
Mut22	30	pMcTs100	T9G/G31D/R32L	614979	614980	pMcTs89
Mut23	31	pMcTs104	T9G	614973	614974	pMeJi9
Mut24	32	pMcTs105	T9G/G31D	614973	614974	pMcTs98
Mut25	33	pMcTs114	G31D/R32L	614470	614471	pMcTs100

Example 3: Expression of *E. coli* ydfG 3-HPDH gene variants in MG1655 cells

[0146] Electrocompetent MG1655 cells were transformed with the resulting cloning plasmids from Example 2 (or controls pMeJi9 or pTrc99A) according to the procedure described in Sheen, J. (1989). "High-Efficiency Transformation by Electroporation." Current Protocols in Molecular Biology. 1.8.4. After a recovery period, two 100 µl aliquots from the transformation reaction were plated onto 150 mm 2XYT plates

supplemented with 100 µg of ampicillin per ml and incubated overnight at 37°C. For each transformation, plasmid DNA of a selected recombinant clone was prepared using a BIOROBOT® 9600 workstation and analyzed by sequencing. The selected clone was then inoculated into a culture of 3 ml of LB media supplemented with 100 µg of ampicillin and incubated overnight at 37°C with shaking. 250 µl of the overnight culture was added to 25 ml of LB media supplemented with 100 µg of ampicillin per ml in a 125 ml baffled shake flask and grown to OD₆₀₀~0.6 before adding 0.5mM IPTG to induce expression from the plasmid. After 1 hour incubation with IPTG the culture was collected by centrifugation and submitted for enzyme assays, as described below. Samples from the cultures were also collected for SDS-PAGE analysis on an 8-16% Bio-Rad Criterion stain-free Tris-HCl gel (Bio-Rad Laboratories, Inc., Hercules, CA, USA).

Example 4: Cofactor specificity of cells expressing *E. coli* ydfG 3-HPDH gene variants

[0147] Cultures from Example 3 were harvested by centrifugation (15,000 x *g* at 4°C for 10 min) and stored at -80°C. Cells were thawed on ice and the pellet was resuspended in Phosphate Buffered Saline (PBS; NaCl, 137 mM; KCl, 2.7 mM; Na₂HPO₄, 10 mM; KH₂PO₄, 1.76 mM) at pH 7.4 containing one tablet of Roche Complete Mini proteases inhibitor cocktail (Roche, Basel) per 10 mL of buffer. Cells were washed three times, and then resuspended in PBS plus protease inhibitor supplemented with lysozyme (Sigma-Aldrich, Saint-Louis, Mo) at a concentration of 2 mg/mL. Cells were then incubated on ice for 30 minutes to allow release of cytoplasmic content, and membrane debris was collected by centrifugation (15,000 x *g* at 4°C for 30 min). The supernatant containing the crude extract (CCE) was transferred to a new tube and kept on ice until further use. CCE protein was quantitated using a Pierce BCA protein detection kit (Thermo Fisher scientific, Rockford, IL, USA) using BSA as a standard by following the manufacturer recommendations. The indicated variants were assayed from the CCE using one or both of the protocols described below.

[0148] A reverse serine dehydrogenase activity assay was conducted with either NADP⁺ or NAD⁺ cofactor by measuring the appearance over time of the associated reduced cofactor at 340 nm. The assay was performed in a 96 well micro-plate, and the final volume was 300 µL. The reaction was started by adding 30 µL of CCE (*supra*) into 270 µL of assay buffer (100 mM Tris pH 8.0, 10 mM NaHCO₃, 5 mM MgCl₂, 400 mM L-serine and 2 mM of either NAD⁺ or NADP⁺). Absorbance at 340 nm was followed on a micro-plate reader (Spectra Max 340PC, Molecular Devices LLC, Sunnyvale, CA, USA) for 10 minutes at room temperature (~25°C). One unit was defined as the amount of enzyme necessary to produce 1 µmol of either NADH or NADPH in one minute in the presence of L-serine at pH 8.0, 25°C.

[0149] The results using the serine dehydrogenase assay (see Table 3) show increased specificity of NAD(H) over to NADP(H) for certain dehydrogenase variants compared to the parent *E. coli* ydfG gene product (expressed from pMeJi9) and a control lacking a ydfG gene product (blank expression vector pTrc99A).

Table 3.

Name	SEQ ID	Cloning Plasmid Name	Serine DeH SA (uts/mg prot)		NAD ⁺ /NADP ⁺
			NADP ⁺	NAD ⁺	
Control (ydfG)	1	pMeJi9	12.74	0.35	0.03
Control	--	pTrc99A	0.35	1.01	2.88
Mut1	7	pMcTs68	0.75	0.58	0.78
Mut2	8	pMcTs69	0.98	2.69	2.76
Mut3	9	pMcTs70	0.28	0.34	1.21

Name	SEQ ID	Cloning Plasmid Name	Serine DeH SA (uts/mg prot)		NAD ⁺ /NADP ⁺
			NADP ⁺	NAD ⁺	
Mut4	10	pMcTs71	0.88	9.41	10.70
Mut5	11	pMcTs72	0.74	0.49	0.66
Mut6	12	pMcTs73	0.80	17.48	21.85
Mut7	13	pMcTs74	0.47	0.92	1.97
Mut8	14	pMcTs75	0.97	1.18	1.22

[0150] A forward malonate semi-aldehyde reductase assay was conducted by measuring the disappearance of either NADH or NADPH over time at 340 nm. Malonate semi-aldehyde was synthesized in-house according to the protocol developed by Yamada and Jacoby (1960) "Direct conversion of malonic semialdehyde to acetyl-coenzyme A", J. Biol. Chem., 235(3): 589-594. The assay was performed in a 96 well micro-plate, and the final volume was 200 μ L. The reaction was started by adding 30 μ L of CCE (*supra*) into 170 μ L of assay buffer (2 mM malonate semialdehyde, 100 mM Tris pH 8.0 and 0.5 mM either NADH or NADPH). Absorbance at 340 nm was followed on a micro-plate reader (Spectra Max 340PC, Molecular Devices LLC) for 10 minutes at room temperature ($\sim 25^{\circ}\text{C}$). One unit was defined as the amount of enzyme necessary to oxidize 1 μ mol of either NADH or NADPH in one minute in the presence of malonate semialdehyde at pH 8.0, 25°C .

[0151] The results using the malonate semi-aldehyde reductase assay (see Table 4) show increased specificity of NAD(H) over to NADP(H) for certain 3-HPDH variants compared to the parent *E. coli* ydfG gene product (expressed from pMeJi9) and a control lacking a ydfG gene product (blank expression vector pTrc99A).

Table 4.

Name	SEQ ID	Cloning Plasmid Name	Serine DeH SA (uts/mg prot)		NADH/NADPH
			NADPH	NADH	
Control	--	pTrc99A	4.30	2.75	0.64
Control (ydfG)	1	pMeJi9	5174.76	0.00	0.00
Mut1	7	pMcTs68	2.18	0.07	0.03
Mut2	8	pMcTs69	4.67	0.00	0.00
Mut3	9	pMcTs70	0.00	0.00	0.00
Mut4	10	pMcTs71	2.75	2.87	1.04
Mut5	11	pMcTs72	4.36	0.00	0.00
Mut6	12	pMcTs73	22.26	39.25	1.76
Mut7	13	pMcTs74	0.75	0.00	0.00
Mut8	14	pMcTs75	3.29	0.00	0.00
Mut9	17	pMcTs79	3.29	0.00	0.00
Mut10	18	pMcTs80	2.63	0.00	0.00
Mut11	19	pMcTs81	1.29	0.00	0.00
Mut12	20	pMcTs82	1.40	31.98	22.84
Mut13	21	pMcTs83	2.16	14.23	6.60

Name	SEQ ID	Cloning Plasmid Name	Serine DeH SA (uts/mg prot)		NADH/NADPH
			NADPH	NADH	
Mut14	22	pMcTs84	2.00	9.41	4.71
Mut15	23	pMcTs85	2.30	20.66	9.00
Mut16	24	pMcTs89	0.44	24.33	54.83
Mut17	25	pMcTs86	0.80	27.45	34.17
Mut18	26	pMcTs87	1.04	22.34	21.47
Mut19	27	pMcTs88	2.27	15.15	6.67
Mut20	28	pMcTs98	0.00	0.00	0.00
Mut21	29	pMcTs99	0.41	5.05	12.45
Mut22	30	pMcTs100	1.55	71.91	46.31
Mut23	31	pMcTs104	288.03	0.00	0.00
Mut24	32	pMcTs105	0.00	10.56	-
Mut25	33	pMcTs114	2.56	0.00	0.00

Example 5: Construction of an expression vector for the *I. orientalis* YMR226c 3-HPDH gene

[0152] The plasmid pMBin190 (WO2012/074818) contains the *I. orientalis* YMR226c nucleotide sequence encoding the 3-HPDH of SEQ ID NO: 4 flanked by NheI/PacI sites. The pMBin190 plasmid was digested with NheI and PacI, gel isolated and purified using Qiagen Gel Extraction kit (Qiagen, Inc.) and the 827bp fragment was ligated into a 7942bp fragment of pMIBa107 (WO2012/074818) digested with XbaI and PacI that was gel isolated and purified using the Qiagen Gel Extraction kit. Cloning of the DNA fragment containing *I. orientalis* YMR226c polynucleotide into pMIBa107 was performed using T4 DNA ligase (New England Biolabs). The reaction mixture contained 1X T4 DNA ligase buffer, 1 µl T4 DNA ligase, 1 µl of the pMIBa107 XbaI/PacI digested DNA fragment above, and 5 µl of the YMR226c NheI/PacI digested product above in total volume of 10µl. The reaction mixture was incubated at room temperature for at least 1 hour and subsequently used to transform One Shot TOP10 cells (Invitrogen) according to manufacturer's instructions. After a recovery period, two 100 µl aliquots from the transformation mixture were plated onto 150 mm 2XYT plates supplemented with 100 µg of ampicillin per ml and incubated overnight at 37°C. Recombinant colonies of the transformations were each inoculated into 3 ml of LB medium supplemented with ampicillin (100 µg/ml). Plasmid DNA was prepared from these cultures using a BIOROBOT® 9600 workstation (Qiagen, Inc.) and subjected restriction digest checks. The plasmid DNA from one clone having the correct restriction digest pattern was further subjected to sequence analysis and designated pMBin200.

[0153] The *I. orientalis* YMR226c coding sequence was amplified by PCR using two synthetic oligonucleotide primers designed to generate an EcoRI restriction site at the 5' end and a XmaI restriction site at the 3' end for integration into pTrc99A (*supra*).

[0154] Twenty nanograms of pMBin200 plasmid DNA was used as a template in a PCR reaction further containing 1X Phusion HF buffer, fifty picomoles each of primers 614967 and 614968, 0.2mM each of dATP, dGTP, dCTP, and dTTP, and 2 units Phusion® Hot Start High-Fidelity DNA Polymerase (Finnzymes, Vantaa, Finland) in a final volume of 50 µl. The amplification reaction was performed in an EPPENDORF® MASTERCYCLER® 5333 (Eppendorf Scientific, Inc.) programmed for one cycle at 95°C for 3 minutes; and 30 cycles each at 95°C for 30 seconds, 56.5°C for 30 seconds, and 72°C for 1 minute. After the 30 cycles,

the reaction was incubated at 72°C for 5 minutes and then cooled at 10°C until further processed.

[0155] The 831bp PCR fragment from the PCR reaction mixture was subjected to 1% TBE-agarose gel electrophoresis with ethidium bromide in TBE buffer and the PCR product was cut out of the gel and purified using the NucleoSpin Extract II kit (Macherey-Nagel). The purified fragment was digested with EcoRI and XmaI (New England Biolabs) and the plasmid pTrc99A was digested with EcoRI and XmaI, and the resulting fragments separated by 1% TBE-agarose gel electrophoresis followed by visualization with a DARK READER™ (Clare Chemical Research). The desired 4.16 kb fragment of pTrc99A and 819bp YMR226c fragment was excised from the gel with a disposable razor blade and purified using a NucleoSpin Extract II kit (Macherey-Nagel).

[0156] Cloning of the DNA fragment containing *I. orientalis* YMR226c coding sequence into pTrc99A was performed using T4 DNA ligase (New England Biolabs). The reaction mixture contained 1X T4 DNA ligase buffer, 1 µl T4 DNA ligase, 1 µl of the pTrc99A EcoRI/XmaI digested DNA fragment above, and 15 µl of the YMR226c EcoRI/XmaI digested PCR product above in total volume of 20µl. The reaction mixture was incubated at room temperature for 1 hour and subsequently used to transform Solo Pack Gold supercompetent cells (Agilent Technologies, Inc.) according to manufacturer's instructions. After a recovery period, two 100 µl aliquots from the transformation mixture were plated onto 150 mm 2XYT plates supplemented with 100 µg of ampicillin per ml and incubated overnight at 37°C. Recombinant colonies of the transformations were each inoculated into 3 ml of LB medium supplemented with ampicillin (100 µg/ml).

[0157] Plasmid DNA was prepared from these cultures using a BIOROBOT® 9600 workstation (Qiagen, Inc.) and analyzed by 1% TBE-agarose gel electrophoresis following EcoRI/XmaI digestion. The plasmid DNA from one clone designated pMcTs103 and having the correct restriction digest pattern was further subjected to sequence analysis to confirm integration of the correct YMR226c coding sequence. From sequencing it was determined that this YMR226c coding sequence differed from the expected genomic sequence by 1 base pair.

[0158] To correct the 1 base pair mutation in pMcTs103 site directed mutagenesis was performed using pMcTs103 as the template DNA in a PCR reaction using QuikChange II XL Site-Directed Mutagenesis Kit (Agilent Technologies, Inc.) as described *supra* using primers 615428 and 614429.

[0159] Recombinant colonies of the transformations were each inoculated into 3 ml of LB medium supplemented with ampicillin (100 µg/ml). Plasmid DNA was prepared from these cultures using a BIOROBOT® 9600 workstation (Qiagen, Inc.) and subjected to sequence analysis to confirm the site directed mutation in the YMR226c coding sequence. A clone with the correct sequence based on sequencing was named pMcTs107.

Example 6: Construction of *I. orientalis* YMR226c 3-HPDH gene variants

[0160] Based on the findings of shown in Example 4, additional 3-HPDH gene variants were constructed using the parent *I. orientalis* YMR226c homolog (SEQ ID NO: 4).

[0161] Site-directed mutagenesis was performed using pMcTs107 (*supra*) as the template DNA in a PCR reaction using a QuikChange II XL Site-Directed Mutagenesis Kit (Agilent Technologies, Inc.) as described *supra* using primers 615006 and 615007, designed to make amino acid substitutions of aspartic acid and leucine at positions 45 and 46 of SEQ ID NO: 4, respectively (corresponding to positions 31 and 32 of SEQ ID NO: 2) resulting in the variant mut26 (SEQ ID NO: 80).

[0162] Recombinant colonies of the transformations were subjected to sequence analysis to confirm the site directed substitutions in the YMR226c coding sequence. A clone with the correct sequence encoding the variant mut26 (SEQ ID NO: 80) was named pMcTs110.

[0163] Site-directed mutagenesis was performed on pMcTs110 in a PCR reaction using a QuikChange II XL Site-Directed Mutagenesis Kit (Agilent Technologies, Inc.) as described *supra* using primers 615004 and 615005, designed to introduce an amino acid substitution of glycine at position 20 of SEQ ID NO: 80 (corresponding to position 9 of SEQ ID NO: 2) resulting in the variant mut27 (SEQ ID NO: 81).

[0164] Recombinant colonies of the transformations were subjected to sequence analysis to confirm the site directed substitutions in the YMR226c coding sequence and a clone with the correct sequence encoding the variant mut27 (SEQ ID NO: 81) was named pMcTs112.

Example 7: Cofactor specificity of cells expressing *I. orientalis* YMR226c 3-HPDH gene variants

[0165] The *I. orientalis* gene variants from Example 6 were expressed in MG1655 cells and the 3-HPDH cofactor specificity was measured using the malonate semi-aldehyde reductase assay described *supra*. Results are shown below in Table 5. The mut27 *I. orientalis* YMR226c 3-HPDH variant expressed from pMcTs112 (SEQ ID NO: 81) showed increased specificity for NAD(H) over to NADP(H) compared to the parent *I. orientalis* YMR226c gene product expressed from pMcTs107 (SEQ ID NO: 4) and a control lacking a YMR226c gene product (blank expression vector pTrc99A).

Table 5.

Name	SEQ ID	Cloning Plasmid Name	Serine DeH SA (uts/mg prot) pH 6		Serine DeH SA (uts/mg prot) pH 8	
			NADPH	NADH	NADPH	NADH
Control	--	pTrc99A	53.32	8.46	7.79	6.96
<i>E. coli</i> ydfG (wt)	1	pMeJi9	12137.66	10.25	737.62	8.73
mut22	30	pMcTs100	36.08	121.98	7.37	38.01
<i>I. orientalis</i> YMR226c (wt)	4	pMcTs107	399.52	8.20	131.39	5.21
mut27	81	pMcTs112	58.02	399.06	5.38	18.76

Example 8: Construction of an expression vector for integration of the *I. orientalis* YMR226c 3-HPDH gene at the *I. orientalis* adh9091 locus

[0166] The *I. orientalis* YMR226 3-HPDH coding sequence was amplified from pMcTs107 (*supra*) with primers designed to add flanking 5' NheI and 3' PaeI restriction sites. Fifty nanograms of pMcTs107 plasmid DNA was used as a template in a PCR reaction further containing 1X Phusion HF buffer, fifty picomoles each of primers 615485 and 615486, 0.2mM each of dATP, dGTP, dCTP, and dTTP, and 2 units Phusion® Hot Start High-Fidelity DNA Polymerase (Finnzymes) in a final volume of 50 µl. The amplification reaction was performed in an EPPENDORF® MASTERCYCLER® 5333 (Eppendorf Scientific, Inc.) programmed for one cycle at 95°C for 3 minutes; and 30 cycles each at 95°C for 30 seconds, 55°C for 30 seconds, and 72°C for 1 minute. After the 30 cycles, the reaction was incubated at 72°C for 5 minutes and then cooled at 10°C until further processed.

[0167] The 837bp PCR fragment from PCR reaction mixture was subjected to 1% TBE-agarose gel electrophoresis with ethidium bromide in TBE buffer and the PCR product was excised from the gel and purified using the NucleoSpin Extract II kit (Macherey-Nagel). The purified fragment was digested with NheI and PacI (New England Biolabs) and the plasmid pMBin204 (WO2012/074818) was digested with XbaI and PacI, and the resulting fragments separated by 1% TBE-agarose gel electrophoresis followed by visualization with a DARK READER™ (Clare Chemical Research). The desired 8.4 kb fragment of pMBin204 and 827bp YMR226c fragment was excised from the gel and purified using a NucleoSpin Extract II kit (Macherey-Nagel).

[0168] Cloning of the DNA fragment containing the coding sequence of the *I. orientalis* YMR226c 3-HPDH (SEQ ID NO: 4) into pMBin204 was performed using T4 DNA ligase (New England Biolabs). The reaction mixture contained 1X T4 DNA ligase buffer, 1 µl T4 DNA ligase, 1 µl of the pMBin204 XbaI and PacI digested DNA fragment above, and 5 µl of the YMR226c NheI and PacI digested PCR product above in total volume of 20 µl. The reaction mixture was incubated at room temperature for 1 hour and subsequently used to transform One Shot TOP10 cells (Invitrogen) according to manufacturer's instructions. After a recovery period, two 100 µl aliquots from the transformation mixture were plated onto 150 mm 2XYT plates supplemented with 100 µg of ampicillin per ml and incubated overnight at 37°C. Recombinant colonies of the transformations were each inoculated into 3 ml of LB medium supplemented with ampicillin (100 µg/ml).

[0169] Plasmid DNA was prepared from these cultures using a BIOROBOT® 9600 workstation (Qiagen, Inc.) and analyzed by restriction digestion. The plasmid DNA from one clone having the correct restriction digest pattern was further subjected to sequence analysis to confirm the correct YMR226c coding sequence was designated pMcTs108 (Figure 5).

Example 9: Construction of an expression vector for integration of the *I. orientalis* YMR226c 3-HPDH gene variants at the *I. orientalis* adh9091 locus

[0170] The coding sequence for the *I. orientalis* YMR226 variant mut27 (SEQ ID NO: 81) was amplified from pMcTs112 (*supra*) with primers designed to add flanking 5' NheI and 3' PacI restriction sites. Fifty nanograms of pMcTs112 plasmid DNA was used as a template in a PCR reaction further containing 1X Phusion HF buffer, fifty picomoles each of primers 615485 and 615486, 0.2mM each of dATP, dGTP, dCTP, and dTTP, and 2 units Phusion® Hot Start High-Fidelity DNA Polymerase (Finnzymes) in a final volume of 50 µl. The amplification reaction was performed in an EPPENDORF® MASTERCYCLER® 5333 (Eppendorf Scientific, Inc.) programmed for one cycle at 95°C for 3 minutes; and 30 cycles each at 95°C for 30 seconds, 55°C for 30 seconds, and 72°C for 1 minute. After the 30 cycles, the reaction was incubated at 72°C for 5 minutes and then cooled at 10°C until further processed.

[0171] The 837bp PCR fragment from PCR reaction mixture was subjected to 1% TBE-agarose gel electrophoresis with ethidium bromide in TBE buffer and the PCR product was cut out of the gel and purified using the NucleoSpin Extract II kit (Macherey-Nagel). The purified fragment was digested with NheI and PacI (New England Biolabs) and the plasmid pMBin204 was digested with XbaI and PacI, and the resulting fragments separated by 1% TBE-agarose gel electrophoresis followed by visualization with a DARK READER™ (Clare Chemical Research). The desired 8.4 kb fragment of pMBin204 and 827bp YMR226c variant fragment was excised from the gel and purified using a NucleoSpin Extract II kit (Macherey-Nagel).

[0172] Cloning of the DNA fragment containing the coding sequence for the *I. orientalis* YMR226c 3-HPDH variant mut27 into pMBin204 was performed using T4 DNA ligase (New England Biolabs). The reaction mixture contained 1X T4 DNA ligase buffer, 1 µl T4 DNA ligase, 1 µl of the pMBin204 XbaI and PacI digested DNA fragment above, and 10 µl of the YMR226c variant NheI and PacI digested PCR product

above in total volume of 20 µl. The reaction mixture was incubated at room temperature for 1 hour and subsequently used to transform One Shot TOP10 cells (Invitrogen) according to manufacturer's instructions. After a recovery period, two 100 µl aliquots from the transformation mixture were plated onto 150 mm 2XYT plates supplemented with 100 µg of ampicillin per ml and incubated overnight at 37°C. Recombinant colonies of the transformations were each inoculated into 3 ml of LB medium supplemented with ampicillin (100 µg/ml).

[0173] Plasmid DNA was prepared from these cultures using a BIOROBOT® 9600 workstation (Qiagen, Inc.) and analyzed by restriction digestion. The plasmid DNA from one clone having the correct restriction digest pattern was further subjected to sequence analysis to confirm the correct YMR226c coding sequence and designated pMcTs116 (Figure 6).

Example 10: Construction of an expression vector for integration of the *E. coli* ydfG 3-HPDH gene at the *I. orientalis* adh9091 locus

[0174] The coding sequence for the *E. coli* ydfG 3-HPDH was codon-optimized for *I. orientalis* DNA, flanked by 5' XbaI site and 3' PaeI restriction sites, and provided by GeneArt in a plasmid designated p1045168 (Figure 7). Plasmids p1045168 and pMBin204 (WO2012/074818) were individually digested with XbaI and PaeI and the resulting fragments separated by 1% TBE-agarose gel electrophoresis and visualized with a DARK READER™ (Clare Chemical Research). The desired 8.4 kb fragment of pMBin204 and 761bp *E. coli* ydfG fragment was excised from the gel and purified using a NucleoSpin Extract II kit (Macherey-Nagel).

[0175] Cloning of the DNA fragment containing the coding sequence of the *E. coli* ydfG 3-HPDH (SEQ ID NO: 2) into pMBin204 was performed using T4 DNA ligase (New England Biolabs). The reaction mixture contained 1X T4 DNA ligase buffer, 1 µl T4 DNA ligase, 1 µl of the pMBin204 XbaI and PaeI digested DNA fragment above, and 10 µl of the *E. coli* ydfG product in total volume of 20µl. The reaction mixture was incubated at room temperature for 1 hour and subsequently used to transform One Shot TOP10 cells (Invitrogen) according to manufacturer's instructions. After a recovery period, two 100 µl aliquots from the transformation mixture were plated onto 150 mm 2XYT plates supplemented with 100 µg of ampicillin per ml and incubated overnight at 37°C. Recombinant colonies of the transformations were each inoculated into 3 ml of LB medium supplemented with ampicillin (100 µg/ml).

[0176] Plasmid DNA was prepared from these cultures using a BIOROBOT® 9600 workstation (Qiagen, Inc.) and analyzed by restriction digestion. The plasmid DNA from one clone having the correct restriction digest pattern was designated pMcTs77 (Figure 8).

Example 11: Construction of an expression vectors for integration of the *E. coli* ydfG 3-HPDH gene variants at the *I. orientalis* adh9091 locus

[0177] The coding sequence for the *E. coli* ydfG 3-HPDH variant mut6 (SEQ ID NO: 12) was codon-optimized for *I. orientalis* DNA, flanked by 5' XbaI site and 3' PaeI restriction sites, and provided by GeneArt in a plasmid designated p11AAT5WP (Figure 9). Fifty nanograms of p11AAT5WP DNA was used as a template in a PCR reaction further containing 1X Expand buffer, fifty picomoles each of primers 614697 and 614698, 0.2mM each of dATP, dGTP, dCTP, and dTTP, and 2.6 units Expand High Fidelity Polymerase (Roche) in a final volume of 50 µl. The amplification reaction was performed in an EPPENDORF® MASTERCYCLER®5333 (Eppendorf Scientific, Inc.) programmed for one cycle at 95°C for 2 minutes; and 30 cycles each at 95°C for 30 seconds, 55°C for 30 seconds, and 72°C for 1 minute. After the 30 cycles, the

reaction was incubated at 72°C for 5 minutes and then cooled at 10°C until further processed.

[0178] The 783bp PCR fragment was digested with XbaI and PaeI and cloned into the 8.4kb fragment of pMBin204 (*supra*) also digested with XbaI and PaeI. Recombinant clones were screened by restriction digest and sequencing and a clone with the correct sequence was designated pMcTs78 (Figure 10).

[0179] Plasmid p1045168 (*supra*; see also Figure 7) was subjected to site directed mutagenesis using primers 615892 and 615893 as described *supra* to change the coding sequence for the wild-type *E. coli* ydfG 3-HPDH (SEQ ID NO: 2) into the coding sequence for the *E. coli* ydfG 3-HPDH variant mut25 (SEQ ID NO: 33) which contains the substitutions G31D and R32L. Recombinant colonies of the transformations were sequenced and a clone encoding the 3-HPDH with the correct amino acid changes was named pMcTs111.

[0180] Plasmid pMcTs111 was subjected to site directed mutagenesis using primers 615890 and 615891 as described *supra* to change the coding sequence for the *E. coli* ydfG 3-HPDH variant mut25 (SEQ ID NO: 33) into the coding sequence for the *E. coli* ydfG 3-HPDH variant mut22 (SEQ ID NO: 30) which contains the an additional T9G substitution. Recombinant colonies of the transformations were sequenced and a clone encoding the 3-HPDH with the correct amino acid changes was named pMcTs111. Recombinant colonies of the transformations were sequenced and a clone encoding the 3-HPDH with the correct amino acid changes was named pMcTs113.

[0181] The coding sequence for the *E. coli* ydfG 3-HPDH variant mut22 (SEQ ID NO: 30) was cloned into pMBin204 by digesting pMcTs113 with XbaI and PaeI and ligating the resulting 761bp fragment of pMcTs113 into the resulting 8.4kbp fragment of pMBin204 also digested with XbaI and PaeI as described *supra*. Recombinant clones were screened by restriction digest and sequencing, and a clone with the correct sequence was designated pMcTs115 (Figure 11).

Example 12: Construction of an expression vectors for integration of the *P. putida* mmsB 3-HPDH at the *I. orientalis* adh9091 locus

[0182] The coding sequence for the *P. putida* mmsB 3-HPDH (SEQ ID NO: 82) was codon-optimized for *I. orientalis* DNA, flanked by 5' XbaI site and 3' PaeI restriction sites, and provided by GeneArt in a plasmid designated p11AA2GJP (Figure 12). Plasmid p11AA2GJP was digested with XbaI and PaeI and the 898bp fragment was cloned into the 8.4kb fragment of pMBin204 also digested with XbaI and PaeI as described *supra*. Several recombinant clones were screened by restriction digest and sequenced. One clone with the correct sequence was designated pMcTs102 (Figure 13).

Example 13: Construction of host strains containing an active 3-HP pathway and expressing 3-HPDH at the *I. orientalis* adh9091 locus

[0183] Approximately 10 µg each of each integration construct pMcTs77, pMcTs78, pMcTs102, pMcTs115 *supra* was individually digested with *Apal* and *KpnI* and separated by gel electrophoresis on a 1% agarose gel using 89 mM Tris base-89 mM Boric Acid-2 mM disodium EDTA (TBE) buffer. Approximately 10 µg each of integration constructs pMcTs108 and pMcTs116 was digested with *Apal* and *SacI* and separated by gel electrophoresis on a 1% agarose gel using TBE buffer. Fragments of approximately 5348 bp for pMcTs77, pMcTs78, and pMcTs115; 5485bp for pMcTs102; and 5408bp for pMcTs108 and pMcTs116 were excised and extracted using the QIAquick gel extraction kit (Qiagen, Inc.) according to the manufacturer's

instructions. The linear constructs from plasmids pMcTs77, pMcTs78, pMcTs102, pMcTs115, pMcTs116, pMcTs108 were transformed into strain McTs259 (containing an active 3-HP pathway but having a deletion to the native *I. orientalis* YMR226c 3-HPDH gene; see WO2012/074818). Several single isolates from each transformation were screened for the site of integration as well as confirming that the other loci were still modified. The integration at adh9091 was confirmed by PCR using Phire® Plant Direct PCR kit (Finnzymes) according to the manufacturer's instructions with primers 614627+612909 and 612908+614626. The PCR product using primers 612908+614626 was approximately 1.97kb for pMcTs77, pMcTs78, pMcTs102, pMcTs115, pMcTs116, and pMcTs108 integrants. The PCR product using primers 614627+612909 was approximately 3.4kb for pMcTs102 integrants and approximately 3.3kb for pMcTs77, pMcTs78, pMcTs102, pMcTs115, pMcTs116, pMcTs108 integrants. The integrity of the existing adh1202 locus and YMR226c locus was verified using primer sets 611245+612794 and 611815+612795 for adh1202 locus, and primer set 613034+613241 for YMR226c locus. A transformant with the correct size bands for the PCRs was designed as show below in Table 6.

Table 6.

Plasmid	3-HPDH gene	Resulting host strain
pMcTs77	<i>E. coli</i> ydfG (wt)	McTs263
pMcTs78	<i>E. coli</i> mut6	McTs265
pMcTs102	<i>P. putida</i> mmsB	McTs276
pMcTs115	<i>E. coli</i> mut22	ShTh100
pMcTs116	<i>I. orientalis</i> mut27	ShTh101
pMcTs108	<i>I. orientalis</i> YMR226c (wt)	MBin556

Example 14: 3-HP production from host strains containing an active 3-HP pathway and expressing 3-HPDH at the *I. orientalis* adh9091 locus

[0184] Strains McTs263, McTs265, McTs276, ShTh100, ShTh101, and MBin556 *supra* were grown in shake flasks and samples were analyzed for cofactor specificity (as described *supra*) and 3-HP production as described in WO2012/074818. Control strains MeJi412 (containing an active 3-HP pathway including the native *I. orientalis* YMR226c 3-HPDH gene) and McTs244 (containing an active 3-HP pathway but having a deletion to the native *I. orientalis* YMR226c 3-HPDH gene) described in WO2012/074818, were also analyzed for 3-HPDH activity and 3-HP production. The results in table 7 show that deletion of the *I. orientalis* YMR226c gene results in no detectable 3-HP production and that 3-HP production can be restored using one copy of a gene encoding a 3-HPDH that has increased specificity for NAD(H).

Table 7.

Strain	3-HPDH gene	3HP (g/L)/ OD600	3HPDH SA, p 8.0		3HPDH SA, pH 6.0	
			NADH	NADPH	NADH	NADPH
McTs263	<i>E. coli</i> ydfG (wt)	0.08	12.10	18.56	47.32	113.91
McTs265	<i>E. coli</i> mut6	0.03	17.95	0.00	90.38	23.15
McTs276	<i>P. putida</i> mmsB	0.06	101.10	0.00	1878.38	193.42
ShTh100	<i>E. coli</i> mut22	0.06	17.25	0.00	72.88	24.156
ShTh101	<i>I. orientalis</i> mut27	0.05	18.30	0.00	221.67	24.58
MBin556	<i>I. orientalis</i> YMR226c (wt)	0.07	13.78	68.86	70.69	292.98

Strain	3-HPDH gene	3HP (g/L)/ OD600	3HPDH SA, p 8.0		3HPDH SA, pH 6.0	
			NADH	NADPH	NADH	NADPH
MeJi412	native	0.14	18.86	47.10	96.33	139.91
McTs244	Deletion of native YMR226c	0.00	12.86	0.00	57.45	19.78

[0185] The invention described and claimed herein is not to be limited in scope by the specific aspects herein disclosed, since these aspects are intended as illustrations of several aspects of the invention. Indeed, various modifications of the invention in addition to those shown and described herein will become apparent to those skilled in the art from the foregoing description. Such modifications are also intended to fall within the scope of the appended claims. In the case of conflict, the present disclosure including definitions will control.

SEQUENCE LISTING

[0186]

<110> Novozymes, Inc. Novozymes A/S Tassone, Monica De Maria, Leonardo

<120> Dehydrogenase variants and polynucleotides encoding same

<130> 12306-WO-PCT

<150> US 61/541,363

<151> 2011-09-30

<160> 82

<170> PatentIn version 3.5

<210> 1

<211> 747

<212> DNA

<213> Escherichia coli

<400> 1

```

atgatcgttt tggtcaccgg tgcaaccgca ggtttcggcg aatgtatcac cagaagattc      60
atccagcagg gtcacaaggt tatcgctacc ggtagaagac aagagagatt gcaagaattg      120
aaggacgagt tgggtgacaa cttgtacatc gctcaattgg acgttagaaa cagagcagct      180
atcgaagaaa tgttggcatc cttgccagct gaatggtgca acatcgacat cttggtcaac      240
aacgctggtt tggcattggg tatggaacca gtcacaagg ctagtgttga ggactgggag      300
accatgatcg acaccaacaa caagggtttg gtctacatga ccagagcagt ttgcctggt      360
atggttgaaa gaaaccacgg tcacatcatc aacatcggtt ccacgctggg ttcctggcca      420
tacgctggcg gtaacgtcta cggtgctacc aaggctttcg ttagacagtt ctcttgaac      480
ttgagaaccg acttgcacgg caccgctgtt agagttaccg acatogaacc aggtttggtt      540
ggtggcaccg aattctccaa cgtcagattc aagggcgacg acggttaaggc tgaaaagacc      600
taccaaaaca cgcgcgtttt gaccccagaa gacgtttcag aggcgtgtttg gtgggtcagt      660

```

accttgccag cacacgtcaa catcaacacc ttggaaatga tgccagtcac ccaatcctac 720
gcagggttga acgttcacag acaataa 747

<210> 2

<211> 248

<212> PRT

<213> Escherichia coli

<400> 2

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

Thr Arg Arg Phe Ile Gln Gln Gly His Lys Val Ile Ala Thr Gly Arg

20 25 30

Arg Gln Glu Arg Leu Gln Glu Leu Lys Asp Glu Leu Gly Asp Asn Leu
35 40 45

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

Leu Ala Ser Leu Pro Ala Glu Trp Cys Asn Ile Asp Ile Leu Val Asn
65 70 75 80

Asn Ala Gly Leu Ala Leu Gly Met Glu Pro Ala His Lys Ala Ser Val
85 90 95

Glu Asp Trp Glu Thr Met Ile Asp Thr Asn Asn Lys Gly Leu Val Tyr
100 105 110

Met Thr Arg Ala Val Leu Pro Gly Met Val Glu Arg Asn His Gly His
115 120 125

Ile Ile Asn Ile Gly Ser Thr Ala Gly Ser Trp Pro Tyr Ala Gly Gly
130 135 140

Asn Val Tyr Gly Ala Thr Lys Ala Phe Val Arg Gln Phe Ser Leu Asn
145 150 155 160

Leu Arg Thr Asp Leu His Gly Thr Ala Val Arg Val Thr Asp Ile Glu
165 170 175

Pro Gly Leu Val Gly Gly Thr Glu Phe Ser Asn Val Arg Phe Lys Gly
180 185 190

Asp Asp Gly Lys Ala Glu Lys Thr Tyr Gln Asn Thr Val Ala Leu Thr
195 200 205

Pro Glu Asp Val Ser Glu Ala Val Trp Trp Val Ser Thr Leu Pro Ala
210 215 220

His Val Asn Ile Asn Thr Leu Glu Met Met Pro Val Thr Gln Ser Tyr
225 230 235 240

Ala Gly Leu Asn Val His Arg Gln
245

<210> 3

<211> 813

<212> DNA

<213> Issatchenkia orientalis

<400> 3

```

atgttttgta atatttccca aagacttgca ggcaagaaca tcctaattac aggtgcgtcc      60
actggtatcg gataccatac agcaaagtat tttgcagaag ctgcaaattg agacttgaag      120
ttggttttgg ctgcaagaag aaaggagaag ctggaggcac taaaggcaga cttgcttgcc      180
aagtatccat ccatcaaagt ccatattgag agtttggatg tctccaaaac ggaaaccatt      240
gcacctttct taaaagggtt acctgaggaa ttttcaattg tcgacgtgtt ggtcaacaat      300
gcaggtaagg cgcttggttt ggatccaatt ggctctgtcg atccaaagga cgtggatgaa      360
atgttccaga ccaatgtttt gggatatgatt caattgaccc agttgggtgt acagcaaattg      420
aaggagagaa actccgggga cattgtccaa ctaggttcag tggctggtag aaaccatac      480
ccaggtggtg gtatctactg tgcttccaag gccgcattga gatcttttac acatgtattg      540
agagaggaat tgattaatac caagattaga gtgattgaaa tcgagcctgg aaatgttgca      600
actgaggaat tttctttgac cagattcaaa ggtgataagt ccaaggccga aaaggtctat      660
gagggaaacc agccattgta tggtagcgat attgcagaat tgattctatt tgcagtttct      720
agacctcaaa acaactgttat tgcagaaaca cttgtttttg ctagtaacca agcttctgct      780
taccatattt tcagaggatc attagataaa tag                                     813

```

<210> 4

<211> 270

<212> PRT

<213> Issatchenkia orientalis

<400> 4

```

Met Phe Gly Asn Ile Ser Gln Arg Leu Ala Gly Lys Asn Ile Leu Ile
 1           5           10          15

Thr Gly Ala Ser Thr Gly Ile Gly Tyr His Thr Ala Lys Tyr Phe Ala
 20          25          30

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

Glu Lys Leu Glu Ala Leu Lys Ala Asp Leu Leu Ala Lys Tyr Pro Ser
 50          55          60

Ile Lys Val His Ile Glu Ser Leu Asp Val Ser Lys Thr Glu Thr Ile
 65          70          75          80

Ala Pro Phe Leu Lys Gly Leu Pro Glu Glu Phe Ser Ile Val Asp Val
 85          90          95

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

Val Asp Pro Lys Asp Val Asp Glu Met Phe Gln Thr Asn Val Leu Gly
115          120          125

Met Ile Gln Leu Thr Gln Leu Val Val Gln Gln Met Lys Glu Arg Asn
130          135          140

Ser Gly Asp Ile Val Gln Leu Gly Ser Val Ala Gly Arg Asn Pro Tyr
145          150          155          160

Pro Gly Gly Gly Ile Tyr Cys Ala Ser Lys Ala Ala Leu Arg Ser Phe
165          170          175

```

100 170 170

Thr His Val Leu Arg Glu Glu Leu Ile Asn Thr Lys Ile Arg Val Ile
180 185 190

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

Phe Lys Gly Asp Lys Ser Lys Ala Glu Lys Val Tyr Glu Gly Thr Glu
210 215 220

Pro Leu Tyr Gly Thr Asp Ile Ala Glu Leu Ile Leu Phe Ala Val Ser
225 230 235 240

Arg Pro Gln Asn Thr Val Ile Ala Glu Thr Leu Val Phe Ala Ser Asn
245 250 255

Gln Ala Ser Ala Tyr His Ile Phe Arg Gly Ser Leu Asp Lys
260 265 270

<210> 5

<211> 804

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 5

atgtcccaag gtagaaaggc agcagaaaga ttggcaaaga agaccgtctt gatcaccggt 60

gcgtccgctg gtatcggtaa ggctaccgcg ttggagtact tggaagcatc caacggtgac 120

atgaagttga tcttggcagc aagaagattg gagaagttgg aagaattgaa gaagaccatc 180

gaccaagaat tcccaaaccg taagggtccac gttgcacaat tggacatcac ccaagcagag 240

aagatcaagc cattcatcga aaacttgcca caagaattca aggacatcga catcttggtc 300

aacaacgctg gtaaggcgtt gggttccgac agagttggtc aaatcgcaac cgaagacatc 360

caagacgtct tcgacaccaa cgtcaccgct ttgatcaaca tcaccaagc tgttttgcca 420

atcttccaag cgaagaactc cggtagacatc gtcaacttgg gttccatcgc tggtagagac 480

gcatacccaa ccgggtccat ctactgcgcc tccaagttcg ctgtcgggtg tttcaccgac 540

tccttgagaa aggaattgat caacaccaag atcagagtca tcttgattgc ccctggtttg 600

gtcgaaacg aattctcctt ggtagatata agaggtaacg aagaacaagc aaagaacgtt 660

tacaaggaca ctacccatt gatggccgac gacgttgacg acttgatcgt ttacgctacc 720

tccagaaagc aaaacaccgt tatcgagac accttgatct tcccaaccaa ccaagcatcc 780

ccacaccaca tcttcagagg ttaa 804

<210> 6

<211> 267

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 6

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> 7

<211> 247

<212> PRT

<213> Escherichia coli

<400> 7

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

Arg Arg Phe Ile Gln Gln Gly His Lys Val Ile Ala Thr Asp Leu Asn
20 25 30

Pro Ala Ala Leu Gln Glu Leu Lys Asp Glu Leu Gly Asp Asn Leu Tyr
35 40 45

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

Ala Ser Leu Pro Ala Glu Trp Cys Asn Ile Asp Ile Leu Val Asn Asn
65 70 75 80

Ala Gly Leu Ala Leu Gly Met Glu Pro Ala His Lys Ala Ser Val Glu
85 90 95

Asp Trp Glu Thr Met Ile Asp Thr Asn Asn Lys Gly Leu Val Tyr Met
100 105 110

Thr Arg Ala Val Leu Pro Gly Met Val Glu Arg Asn His Gly His Ile
115 120 125

Ile Asn Ile Gly Ser Thr Ala Gly Ser Trp Pro Tyr Ala Gly Gly Asn
130 135 140

Val Tyr Gly Ala Thr Lys Ala Phe Val Arg Gln Phe Ser Leu Asn Leu
145 150 155 160

Arg Thr Asp Leu His Gly Thr Ala Val Arg Val Thr Asp Ile Glu Pro
165 170 175

Gly Leu Val Gly Gly Thr Glu Phe Ser Asn Val Arg Phe Lys Gly Asp
180 185 190

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

Glu Asp Val Ser Glu Ala Val Trp Trp Val Ser Thr Leu Pro Ala His
210 215 220

Val Asn Ile Asn Thr Leu Glu Met Met Pro Val Thr Gln Ser Tyr Ala
225 230 235 240

Gly Leu Asn Val His Arg Gln
245

<210> 8

<211> 248

<212> PRT

<213> Escherichia coli

<400> 8

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

Thr Arg Arg Phe Ile Gln Gln Gly His Lys Val Ile Ala Thr Glu Leu
20 25 30

Asn Pro Ala Ala Leu Gln Glu Leu Lys Asp Glu Leu Gly Asp Asn Leu
35 40 45

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

Leu Ala Ser Leu Pro Ala Glu Trp Cys Asn Ile Asp Ile Leu Val Asn
65 70 75 80

Asn Ala Gly Leu Ala Leu Gly Met Glu Pro Ala His Lys Ala Ser Val
85 90 95

Glu Asp Trp Glu Thr Met Ile Asp Thr Asn Asn Lys Gly Leu Val Tyr
100 105 110

Met Thr Arg Ala Val Leu Pro Gly Met Val Glu Arg Asn His Gly His
115 120 125

Ile Ile Asn Ile Gly Ser Thr Ala Gly Ser Trp Pro Tyr Ala Gly Gly
130 135 140

Asn Val Tyr Gly Ala Thr Lys Ala Phe Val Arg Gln Phe Ser Leu Asn
145 150 155 160

Leu Arg Thr Asp Leu His Gly Thr Ala Val Arg Val Thr Asp Ile Glu
165 170 175

Pro Gly Leu Val Gly Gly Thr Glu Phe Ser Asn Val Arg Phe Lys Gly
180 185 190

Asp Asp Gly Lys Ala Glu Lys Thr Tyr Gln Asn Thr Val Ala Leu Thr
195 200 205

Pro Glu Asp Val Ser Glu Ala Val Trp Trp Val Ser Thr Leu Pro Ala
210 215 220

His Val Asn Ile Asn Thr Leu Glu Met Met Pro Val Thr Gln Ser Tyr
225 230 235 240

Ala Gly Leu Asn Val His Arg Gln
245

<210> 9

<211> 247

<212> PRT

<213> Escherichia coli

<400> 9

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

Arg Arg Phe Ile Gln Gln Gly His Lys Val Ile Ala Thr Glu Leu Asn
20 25 30

Pro Ala Ala Leu Gln Glu Leu Lys Asp Glu Leu Gly Asp Asn Leu Tyr
35 40 45

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

Ala Ser Leu Pro Ala Glu Trp Cys Asn Ile Asp Ile Leu Val Asn Asn
65 70 75 80

Ala Gly Leu Ala Leu Gly Met Glu Pro Ala His Lys Ala Ser Val Glu
85 90 95

Asp Trp Glu Thr Met Ile Asp Thr Asn Asn Lys Gly Leu Val Tyr Met
100 105 110

Thr Arg Ala Val Leu Pro Gly Met Val Glu Arg Asn His Gly His Ile
115 120 125

Ile Asn Ile Glv Ser Thr Ala Glv Ser Trp Pro Trv Ala Glv Glv Asn

```

-----
130              135              140
Val Tyr Gly Ala Thr Lys Ala Phe Val Arg Gln Phe Ser Leu Asn Leu
145              150              155              160

Arg Thr Asp Leu His Gly Thr Ala Val Arg Val Thr Asp Ile Glu Pro
165              170              175

Gly Leu Val Gly Gly Thr Glu Phe Ser Asn Val Arg Phe Lys Gly Asp
180              185              190

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

Glu Asp Val Ser Glu Ala Val Trp Trp Val Ser Thr Leu Pro Ala His
210              215              220

Val Asn Ile Asn Thr Leu Glu Met Met Pro Val Thr Gln Ser Tyr Ala
225              230              235              240

Gly Leu Asn Val His Arg Gln
245

<210> 10
<211> 248
<212> PRT
<213> Escherichia coli

<400> 10
Met Ile Val Leu Val Thr Gly Ala Gly Ala Gly Phe Gly Glu Cys Ile
1      5      10      15

Thr Arg Arg Phe Ile Gln Gln Gly His Lys Val Ile Ala Thr Asp Leu
20      25      30

Ser Ala Asp Ala Leu Gln Glu Leu Lys Asp Glu Leu Gly Asp Asn Leu
35      40      45

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

Leu Ala Ser Leu Pro Ala Glu Trp Cys Asn Ile Asp Ile Leu Val Asn
65      70      75      80

Asn Ala Gly Leu Ala Leu Gly Met Glu Pro Ala His Lys Ala Ser Val
85      90      95

Glu Asp Trp Glu Thr Met Ile Asp Thr Asn Asn Lys Gly Leu Val Tyr
100     105     110

Met Thr Arg Ala Val Leu Pro Gly Met Val Glu Arg Asn His Gly His
115     120     125

Ile Ile Asn Ile Gly Ser Thr Ala Gly Ser Trp Pro Tyr Ala Gly Gly
130     135     140

Asn Val Tyr Gly Ala Thr Lys Ala Phe Val Arg Gln Phe Ser Leu Asn
145     150     155     160

Leu Arg Thr Asp Leu His Gly Thr Ala Val Arg Val Thr Asp Ile Glu
165     170     175

```

Pro Gly Leu Val Gly Gly Thr Glu Phe Ser Asn Val Arg Phe Lys Gly
180 185 190

Asp Asp Gly Lys Ala Glu Lys Thr Tyr Gln Asn Thr Val Ala Leu Thr
195 200 205

Pro Glu Asp Val Ser Glu Ala Val Trp Trp Val Ser Thr Leu Pro Ala
210 215 220

His Val Asn Ile Asn Thr Leu Glu Met Met Pro Val Thr Gln Ser Tyr
225 230 235 240

Ala Gly Leu Asn Val His Arg Gln
245

<210> 11

<211> 247

<212> PRT

<213> Escherichia coli

<400> 11

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

Arg Arg Phe Ile Gln Gln Gly His Lys Val Ile Ala Thr Asp Leu Ser
20 25 30

Ala Asp Ala Leu Gln Glu Leu Lys Asp Glu Leu Gly Asp Asn Leu Tyr
35 40 45

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

Ala Ser Leu Pro Ala Glu Trp Cys Asn Ile Asp Ile Leu Val Asn Asn
65 70 75 80

Ala Gly Leu Ala Leu Gly Met Glu Pro Ala His Lys Ala Ser Val Glu
85 90 95

Asp Trp Glu Thr Met Ile Asp Thr Asn Asn Lys Gly Leu Val Tyr Met
100 105 110

Thr Arg Ala Val Leu Pro Gly Met Val Glu Arg Asn His Gly His Ile
115 120 125

Ile Asn Ile Gly Ser Thr Ala Gly Ser Trp Pro Tyr Ala Gly Gly Asn
130 135 140

Val Tyr Gly Ala Thr Lys Ala Phe Val Arg Gln Phe Ser Leu Asn Leu
145 150 155 160

Arg Thr Asp Leu His Gly Thr Ala Val Arg Val Thr Asp Ile Glu Pro
165 170 175

Gly Leu Val Gly Gly Thr Glu Phe Ser Asn Val Arg Phe Lys Gly Asp
180 185 190

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

Glu Asp Val Ser Glu Ala Val Trp Trp Val Ser Thr Leu Pro Ala His
210 215 220

210 215 220
 Val Asn Ile Asn Thr Leu Glu Met Met Pro Val Thr Gln Ser Tyr Ala
 225 230 235 240

 Gly Leu Asn Val His Arg Gln
 245

 <210> 12
 <211> 248
 <212> PRT
 <213> Escherichia coli

 <400> 12
 Met Ile Val Leu Val Thr Gly Ala Gly Ala Gly Phe Gly Glu Cys Ile
 1 5 10 15

 Thr Arg Arg Phe Ile Gln Gln Gly His Lys Val Ile Ala Thr Asp Leu
 20 25 30

 Asn Pro Ala Ala Leu Gln Glu Leu Lys Asp Glu Leu Gly Asp Asn Leu
 35 40 45

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

 Leu Ala Ser Leu Pro Ala Glu Trp Cys Asn Ile Asp Ile Leu Val Asn
 65 70 75 80

 Asn Ala Gly Leu Ala Leu Gly Met Glu Pro Ala His Lys Ala Ser Val
 85 90 95

 Glu Asp Trp Glu Thr Met Ile Asp Thr Asn Asn Lys Gly Leu Val Tyr
 100 105 110

 Met Thr Arg Ala Val Leu Pro Gly Met Val Glu Arg Asn His Gly His
 115 120 125

 Ile Ile Asn Ile Gly Ser Thr Ala Gly Ser Trp Pro Tyr Ala Gly Gly
 130 135 140

 Asn Val Tyr Gly Ala Thr Lys Ala Phe Val Arg Gln Phe Ser Leu Asn
 145 150 155 160

 Leu Arg Thr Asp Leu His Gly Thr Ala Val Arg Val Thr Asp Ile Glu
 165 170 175

 Pro Gly Leu Val Gly Gly Thr Glu Phe Ser Asn Val Arg Phe Lys Gly
 180 185 190

 Asp Asp Gly Lys Ala Glu Lys Thr Tyr Gln Asn Thr Val Ala Leu Thr
 195 200 205

 Pro Glu Asp Val Ser Glu Ala Val Trp Trp Val Ser Thr Leu Pro Ala
 210 215 220

 His Val Asn Ile Asn Thr Leu Glu Met Met Pro Val Thr Gln Ser Tyr
 225 230 235 240

 Ala Gly Leu Asn Val His Arg Gln
 245

<210> 13

<211> 247

<212> PRT

<213> Escherichia coli

<400> 13

```

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

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

Ala Asp Ala Leu Gln Glu Leu Lys Asp Glu Leu Gly Asp Asn Leu Tyr
 35              40              45

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

Ala Ser Leu Pro Ala Glu Trp Cys Asn Ile Asp Ile Leu Val Asn Asn
 65              70              75              80

Ala Gly Leu Ala Leu Gly Met Glu Pro Ala His Lys Ala Ser Val Glu
 85              90              95

Asp Trp Glu Thr Met Ile Asp Thr Asn Asn Lys Gly Leu Val Tyr Met
100              105              110

Thr Arg Ala Val Leu Pro Gly Met Val Glu Arg Asn His Gly His Ile
115              120              125

Ile Asn Ile Gly Ser Thr Ala Gly Ser Trp Pro Tyr Ala Gly Gly Asn
130              135              140

Val Tyr Gly Ala Thr Lys Ala Phe Val Arg Gln Phe Ser Leu Asn Leu
145              150              155              160

Arg Thr Asp Leu His Gly Thr Ala Val Arg Val Thr Asp Ile Glu Pro
165              170              175

Gly Leu Val Gly Gly Thr Glu Phe Ser Asn Val Arg Phe Lys Gly Asp
180              185              190

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

Glu Asp Val Ser Glu Ala Val Trp Trp Val Ser Thr Leu Pro Ala His
210              215              220

Val Asn Ile Asn Thr Leu Glu Met Met Pro Val Thr Gln Ser Tyr Ala
225              230              235              240

Gly Leu Asn Val His Arg Gln
245

```

<210> 14

<211> 248

<212> PRT

<213> Escherichia coli

<400> 14

```

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

```

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

<210> 15

<211> 31

<212> DNA

<213> Escherichia coli

<400> 15

cggaattcat gatcgttta gtaactggag c 31

<210> 16

<211> 29

<212> DNA

<213> Escherichia coli

<400> 16

cgggatcctt actgacggtg gacattcag 29

<210> 17

<211> 248

<212> PRT

<213> Escherichia coli

<400> 17

```

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

Thr Arg Arg Phe Ile Gln Gln Gly His Lys Val Ile Ala Thr Asp Leu
      20             25             30

Asn Pro Ala Ala Leu Gln Glu Leu Lys Asp Glu Leu Gly Asp Asn Leu
      35             40             45

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

Leu Ala Ser Leu Pro Ala Glu Trp Cys Asn Ile Asp Ile Leu Val Asn
 65             70             75             80

Asn Ala Gly Leu Ala Leu Gly Met Glu Pro Ala His Lys Ala Ser Val
      85             90             95

Glu Asp Trp Glu Thr Met Ile Asp Thr Asn Asn Lys Gly Leu Val Tyr
      100             105             110

Met Thr Arg Ala Val Leu Pro Gly Met Val Glu Arg Asn His Gly His
      115             120             125

Ile Ile Asn Ile Gly Ser Thr Ala Gly Ser Trp Pro Tyr Ala Gly Gly
      130             135             140

Asn Val Tyr Gly Ala Thr Lys Ala Phe Val Arg Gln Phe Ser Leu Asn
      145             150             155             160

Leu Arg Thr Asp Leu His Gly Thr Ala Val Arg Val Thr Asp Ile Glu
      165             170             175

Pro Gly Leu Val Gly Gly Thr Glu Phe Ser Asn Val Arg Phe Lys Gly
      180             185             190

Asp Asp Gly Lys Ala Glu Lys Thr Tyr Gln Asn Thr Val Ala Leu Thr
      195             200             205

Pro Glu Asp Val Ser Glu Ala Val Trp Trp Val Ser Thr Leu Pro Ala
      210             215             220

His Val Asn Ile Asn Thr Leu Glu Met Met Pro Val Thr Gln Ser Tyr
      225             230             235             240

Ala Gly Leu Asn Val His Arg Gln
      245

```

<210> 18

<211> 248

<212> PRT

<213> Escherichia coli

<400> 18

```

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

```

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

<210> 19

<211> 248

<212> PRT

<213> Escherichia coli

<400> 19

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

Leu Ala Ser Leu Pro Ala Glu Trp Cys Asn Ile Asp Ile Leu Val Asn
65 70 75 80

Asn Ala Gly Leu Ala Leu Gly Met Glu Pro Ala His Lys Ala Ser Val
85 90 95

Glu Asp Trp Glu Thr Met Ile Asp Thr Asn Asn Lys Gly Leu Val Tyr
100 105 110

Met Thr Arg Ala Val Leu Pro Gly Met Val Glu Arg Asn His Gly His
115 120 125

Ile Ile Asn Ile Gly Ser Thr Ala Gly Ser Trp Pro Tyr Ala Gly Gly
130 135 140

Asn Val Tyr Gly Ala Thr Lys Ala Phe Val Arg Gln Phe Ser Leu Asn
145 150 155 160

Leu Arg Thr Asp Leu His Gly Thr Ala Val Arg Val Thr Asp Ile Glu
165 170 175

Pro Gly Leu Val Gly Gly Thr Glu Phe Ser Asn Val Arg Phe Lys Gly
180 185 190

Asp Asp Gly Lys Ala Glu Lys Thr Tyr Gln Asn Thr Val Ala Leu Thr
195 200 205

Pro Glu Asp Val Ser Glu Ala Val Trp Trp Val Ser Thr Leu Pro Ala
210 215 220

His Val Asn Ile Asn Thr Leu Glu Met Met Pro Val Thr Gln Ser Tyr
225 230 235 240

Ala Gly Leu Asn Val His Arg Gln
245

<210> 20

<211> 248

<212> PRT

<213> Escherichia coli

<400> 20

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

Thr Arg Arg Phe Ile Gln Gln Gly His Lys Val Ile Ala Thr Asp Leu
20 25 30

Arg Pro Ala Ala Leu Gln Glu Leu Lys Asp Glu Leu Gly Asp Asn Leu
35 40 45

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

Leu Ala Ser Leu Pro Ala Glu Trp Cys Asn Ile Asp Ile Leu Val Asn
65 70 75 80

Asn Ala Gly Leu Ala Leu Gly Met Glu Pro Ala His Lys Ala Ser Val
85 90 95

Glu Asp Trp Glu Thr Met Ile Asp Thr Asn Asn Lys Gly Leu Val Tyr
100 105 110

Met Thr Arg Ala Val Leu Pro Gly Met Val Glu Arg Asn His Gly His
115 120 125

Ile Ile Asn Ile Gly Ser Thr Ala Gly Ser Trp Pro Tyr Ala Gly Gly
130 135 140

Asn Val Tyr Gly Ala Thr Lys Ala Phe Val Arg Gln Phe Ser Leu Asn
145 150 155 160

Leu Arg Thr Asp Leu His Gly Thr Ala Val Arg Val Thr Asp Ile Glu
165 170 175

Pro Gly Leu Val Gly Gly Thr Glu Phe Ser Asn Val Arg Phe Lys Gly
180 185 190

Asp Asp Gly Lys Ala Glu Lys Thr Tyr Gln Asn Thr Val Ala Leu Thr
195 200 205

Pro Glu Asp Val Ser Glu Ala Val Trp Trp Val Ser Thr Leu Pro Ala
210 215 220

His Val Asn Ile Asn Thr Leu Glu Met Met Pro Val Thr Gln Ser Tyr
225 230 235 240

Ala Gly Leu Asn Val His Arg Gln
245

<210> 21

<211> 248

<212> PRT

<213> Escherichia coli

<400> 21

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

Thr Arg Arg Phe Ile Gln Gln Gly His Lys Val Ile Ala Thr Asp Leu
20 25 30

Asn Gln Ala Ala Leu Gln Glu Leu Lys Asp Glu Leu Gly Asp Asn Leu
35 40 45

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

Leu Ala Ser Leu Pro Ala Glu Trp Cys Asn Ile Asp Ile Leu Val Asn
65 70 75 80

Asn Ala Gly Leu Ala Leu Gly Met Glu Pro Ala His Lys Ala Ser Val
85 90 95

Glu Asp Trp Glu Thr Met Ile Asp Thr Asn Asn Lys Gly Leu Val Tyr
100 105 110

Met Thr Arg Ala Val Leu Pro Gly Met Val Glu Arg Asn His Gly His
115 120 125

Ile Ile Asn Ile Gly Ser Thr Ala Gly Ser Trp Pro Tyr Ala Gly Gly
130 135 140

Asn Val Tyr Gly Ala Thr Lys Ala Phe Val Arg Gln Phe Ser Leu Asn
145 150 155 160

Leu Arg Thr Asp Leu His Gly Thr Ala Val Arg Val Thr Asp Ile Glu
165 170 175

Pro Gly Leu Val Gly Gly Thr Glu Phe Ser Asn Val Arg Phe Lys Gly
180 185 190

Asp Asp Gly Lys Ala Glu Lys Thr Tyr Gln Asn Thr Val Ala Leu Thr
195 200 205

Pro Glu Asp Val Ser Glu Ala Val Trp Trp Val Ser Thr Leu Pro Ala
210 215 220

His Val Asn Ile Asn Thr Leu Glu Met Met Pro Val Thr Gln Ser Tyr
225 230 235 240

Ala Gly Leu Asn Val His Arg Gln
245

<210> 22

<211> 248

<212> PRT

<213> Escherichia coli

<400> 22

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

Thr Arg Arg Phe Ile Gln Gln Gly His Lys Val Ile Ala Thr Asp Leu
20 25 30

Asn Pro Glu Ala Leu Gln Glu Leu Lys Asp Glu Leu Gly Asp Asn Leu
35 40 45

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

Leu Ala Ser Leu Pro Ala Glu Trp Cys Asn Ile Asp Ile Leu Val Asn
65 70 75 80

Asn Ala Gly Leu Ala Leu Gly Met Glu Pro Ala His Lys Ala Ser Val
85 90 95

Glu Asp Trp Glu Thr Met Ile Asp Thr Asn Asn Lys Gly Leu Val Tyr
100 105 110

Met Thr Arg Ala Val Leu Pro Gly Met Val Glu Arg Asn His Gly His
115 120 125

Ile Ile Asn Ile Gly Ser Thr Ala Gly Ser Trp Pro Tyr Ala Gly Gly
130 135 140

Asn Val Tyr Gly Ala Thr Lys Ala Phe Val Arg Gln Phe Ser Leu Asn
145 150 155 160

Leu Arg Thr Asp Leu His Gly Thr Ala Val Arg Val Thr Asp Ile Glu
165 170 175

Pro Gly Leu Val Gly Gly Thr Glu Phe Ser Asn Val Arg Phe Lys Gly
180 185 190

Asp Asp Gly Lys Ala Glu Lys Thr Tyr Gln Asn Thr Val Ala Leu Thr
195 200 205

Pro Glu Asp Val Ser Glu Ala Val Trp Trp Val Ser Thr Leu Pro Ala
210 215 220

His Val Asn Ile Asn Thr Leu Glu Met Met Pro Val Thr Gln Ser Tyr
225 230 235 240

Ala Gly Leu Asn Val His Arg Gln
245

<210> 23

<211> 248

<212> PRT

<213> Escherichia coli

<400> 23

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

Thr Arg Arg Phe Ile Gln Gln Gly His Lys Val Ile Ala Thr Asp Leu
20 25 30

Asn Pro Ala Arg Leu Gln Glu Leu Lys Asp Glu Leu Gly Asp Asn Leu
35 40 45

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

Leu Ala Ser Leu Pro Ala Glu Trp Cys Asn Ile Asp Ile Leu Val Asn
65 70 75 80

Asn Ala Gly Leu Ala Leu Gly Met Glu Pro Ala His Lys Ala Ser Val
85 90 95

Glu Asp Trp Glu Thr Met Ile Asp Thr Asn Asn Lys Gly Leu Val Tyr
100 105 110

Met Thr Arg Ala Val Leu Pro Gly Met Val Glu Arg Asn His Gly His
115 120 125

Ile Ile Asn Ile Gly Ser Thr Ala Gly Ser Trp Pro Tyr Ala Gly Gly
130 135 140

Asn Val Tyr Gly Ala Thr Lys Ala Phe Val Arg Gln Phe Ser Leu Asn
145 150 155 160

Leu Arg Thr Asp Leu His Gly Thr Ala Val Arg Val Thr Asp Ile Glu
165 170 175

Pro Gly Leu Val Gly Gly Thr Glu Phe Ser Asn Val Arg Phe Lys Gly
180 185 190

Asp Asp Gly Lys Ala Glu Lys Thr Tyr Gln Asn Thr Val Ala Leu Thr

```

195          200          205

Pro Glu Asp Val Ser Glu Ala Val Trp Trp Val Ser Thr Leu Pro Ala
210          215          220

His Val Asn Ile Asn Thr Leu Glu Met Met Pro Val Thr Gln Ser Tyr
225          230          235          240

Ala Gly Leu Asn Val His Arg Gln
245

<210> 24
<211> 248
<212> PRT
<213> Escherichia coli

<400> 24
Met Ile Val Leu Val Thr Gly Ala Gly Ala Gly Phe Gly Glu Cys Ile
1          5          10          15

Thr Arg Arg Phe Ile Gln Gln Gly His Lys Val Ile Ala Thr Asp Leu
20          25          30

Arg Gln Glu Ala Leu Gln Glu Leu Lys Asp Glu Leu Gly Asp Asn Leu
35          40          45

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

Leu Ala Ser Leu Pro Ala Glu Trp Cys Asn Ile Asp Ile Leu Val Asn
65          70          75          80

Asn Ala Gly Leu Ala Leu Gly Met Glu Pro Ala His Lys Ala Ser Val
85          90          95

Glu Asp Trp Glu Thr Met Ile Asp Thr Asn Asn Lys Gly Leu Val Tyr
100         105         110

Met Thr Arg Ala Val Leu Pro Gly Met Val Glu Arg Asn His Gly His
115         120         125

Ile Ile Asn Ile Gly Ser Thr Ala Gly Ser Trp Pro Tyr Ala Gly Gly
130         135         140

Asn Val Tyr Gly Ala Thr Lys Ala Phe Val Arg Gln Phe Ser Leu Asn
145         150         155         160

Leu Arg Thr Asp Leu His Gly Thr Ala Val Arg Val Thr Asp Ile Glu
165         170         175

Pro Gly Leu Val Gly Gly Thr Glu Phe Ser Asn Val Arg Phe Lys Gly
180         185         190

Asp Asp Gly Lys Ala Glu Lys Thr Tyr Gln Asn Thr Val Ala Leu Thr
195         200         205

Pro Glu Asp Val Ser Glu Ala Val Trp Trp Val Ser Thr Leu Pro Ala
210         215         220

His Val Asn Ile Asn Thr Leu Glu Met Met Pro Val Thr Gln Ser Tyr
225         230         235         240

```

Ala Gly Leu Asn Val His Arg Gln
245

<210> 25

<211> 248

<212> PRT

<213> Escherichia coli

<400> 25

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

Thr Arg Arg Phe Ile Gln Gln Gly His Lys Val Ile Ala Thr Asp Leu
20 25 30

Arg Gln Ala Ala Leu Gln Glu Leu Lys Asp Glu Leu Gly Asp Asn Leu
35 40 45

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

Leu Ala Ser Leu Pro Ala Glu Trp Cys Asn Ile Asp Ile Leu Val Asn
65 70 75 80

Asn Ala Gly Leu Ala Leu Gly Met Glu Pro Ala His Lys Ala Ser Val
85 90 95

Glu Asp Trp Glu Thr Met Ile Asp Thr Asn Asn Lys Gly Leu Val Tyr
100 105 110

Met Thr Arg Ala Val Leu Pro Gly Met Val Glu Arg Asn His Gly His
115 120 125

Ile Ile Asn Ile Gly Ser Thr Ala Gly Ser Trp Pro Tyr Ala Gly Gly
130 135 140

Asn Val Tyr Gly Ala Thr Lys Ala Phe Val Arg Gln Phe Ser Leu Asn
145 150 155 160

Leu Arg Thr Asp Leu His Gly Thr Ala Val Arg Val Thr Asp Ile Glu
165 170 175

Pro Gly Leu Val Gly Gly Thr Glu Phe Ser Asn Val Arg Phe Lys Gly
180 185 190

Asp Asp Gly Lys Ala Glu Lys Thr Tyr Gln Asn Thr Val Ala Leu Thr
195 200 205

Pro Glu Asp Val Ser Glu Ala Val Trp Trp Val Ser Thr Leu Pro Ala
210 215 220

His Val Asn Ile Asn Thr Leu Glu Met Met Pro Val Thr Gln Ser Tyr
225 230 235 240

Ala Gly Leu Asn Val His Arg Gln
245

<210> 26

<211> 248

<212> PRT

<213> Escherichia coli

<400> 26

Met Ile Val Leu Val Thr Gly Ala Gly Ala Gly Phe Gly Glu Cys Ile

1 5 10 15

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

<210> 27

<211> 248

<212> PRT

<213> Escherichia coli

<400> 27

Met Ile Val Leu Val Thr Gly Ala Gly Ala Gly Phe Gly Glu Cys Ile
1 5 10 15Thr Arg Arg Phe Ile Gln Gln Gly His Lys Val Ile Ala Thr Asp Leu
20 25 30

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

<210> 28

<211> 248

<212> PRT

<213> Escherichia coli

<400> 28

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

```

Asn Ala Gly Leu Ala Leu Gly Met Glu Pro Ala His Lys Ala Ser Val
      85                      90                      95

Glu Asp Trp Glu Thr Met Ile Asp Thr Asn Asn Lys Gly Leu Val Tyr
      100                      105                      110

Met Thr Arg Ala Val Leu Pro Gly Met Val Glu Arg Asn His Gly His
      115                      120                      125

Ile Ile Asn Ile Gly Ser Thr Ala Gly Ser Trp Pro Tyr Ala Gly Gly
      130                      135                      140

Asn Val Tyr Gly Ala Thr Lys Ala Phe Val Arg Gln Phe Ser Leu Asn
      145                      150                      155                      160

Leu Arg Thr Asp Leu His Gly Thr Ala Val Arg Val Thr Asp Ile Glu
      165                      170                      175

Pro Gly Leu Val Gly Gly Thr Glu Phe Ser Asn Val Arg Phe Lys Gly
      180                      185                      190

Asp Asp Gly Lys Ala Glu Lys Thr Tyr Gln Asn Thr Val Ala Leu Thr
      195                      200                      205

Pro Glu Asp Val Ser Glu Ala Val Trp Trp Val Ser Thr Leu Pro Ala
      210                      215                      220

His Val Asn Ile Asn Thr Leu Glu Met Met Pro Val Thr Gln Ser Tyr
      225                      230                      235                      240

Ala Gly Leu Asn Val His Arg Gln
      245

```

<210> 29

<211> 248

<212> PRT

<213> Escherichia coli

<400> 29

```

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

Thr Arg Arg Phe Ile Gln Gln Gly His Lys Val Ile Ala Thr Asp Arg
      20                      25                      30

Arg Gln Glu Ala Leu Gln Glu Leu Lys Asp Glu Leu Gly Asp Asn Leu
      35                      40                      45

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

Leu Ala Ser Leu Pro Ala Glu Trp Cys Asn Ile Asp Ile Leu Val Asn
      65                      70                      75                      80

Asn Ala Gly Leu Ala Leu Gly Met Glu Pro Ala His Lys Ala Ser Val
      85                      90                      95

Glu Asp Trp Glu Thr Met Ile Asp Thr Asn Asn Lys Gly Leu Val Tyr
      100                      105                      110

Met Thr Arg Ala Val Leu Pro Gly Met Val Glu Arg Asn His Gly His
      115                      120                      125

```

```

110
120
130
Ile Ile Asn Ile Gly Ser Thr Ala Gly Ser Trp Pro Tyr Ala Gly Gly
130          135          140

Asn Val Tyr Gly Ala Thr Lys Ala Phe Val Arg Gln Phe Ser Leu Asn
145          150          155          160

Leu Arg Thr Asp Leu His Gly Thr Ala Val Arg Val Thr Asp Ile Glu
165          170          175

Pro Gly Leu Val Gly Gly Thr Glu Phe Ser Asn Val Arg Phe Lys Gly
180          185          190

Asp Asp Gly Lys Ala Glu Lys Thr Tyr Gln Asn Thr Val Ala Leu Thr
195          200          205

Pro Glu Asp Val Ser Glu Ala Val Trp Trp Val Ser Thr Leu Pro Ala
210          215          220

His Val Asn Ile Asn Thr Leu Glu Met Met Pro Val Thr Gln Ser Tyr
225          230          235          240

Ala Gly Leu Asn Val His Arg Gln
245

```

<210> 30

<211> 248

<212> PRT

<213> Escherichia coli

<400> 30

```

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

Thr Arg Arg Phe Ile Gln Gln Gly His Lys Val Ile Ala Thr Asp Leu
20          25          30

Arg Gln Glu Arg Leu Gln Glu Leu Lys Asp Glu Leu Gly Asp Asn Leu
35          40          45

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

Leu Ala Ser Leu Pro Ala Glu Trp Cys Asn Ile Asp Ile Leu Val Asn
65          70          75          80

Asn Ala Gly Leu Ala Leu Gly Met Glu Pro Ala His Lys Ala Ser Val
85          90          95

Glu Asp Trp Glu Thr Met Ile Asp Thr Asn Asn Lys Gly Leu Val Tyr
100          105          110

Met Thr Arg Ala Val Leu Pro Gly Met Val Glu Arg Asn His Gly His
115          120          125

Ile Ile Asn Ile Gly Ser Thr Ala Gly Ser Trp Pro Tyr Ala Gly Gly
130          135          140

Asn Val Tyr Gly Ala Thr Lys Ala Phe Val Arg Gln Phe Ser Leu Asn
145          150          155          160

Leu Arg Thr Asp Leu His Gly Thr Ala Val Arg Val Thr Asp Ile Glu
165          170          175

```

Pro Gly Leu Val Gly Gly Thr Glu Phe Ser Asn Val Arg Phe Lys Gly
180 185 190

Asp Asp Gly Lys Ala Glu Lys Thr Tyr Gln Asn Thr Val Ala Leu Thr
195 200 205

Pro Glu Asp Val Ser Glu Ala Val Trp Trp Val Ser Thr Leu Pro Ala
210 215 220

His Val Asn Ile Asn Thr Leu Glu Met Met Pro Val Thr Gln Ser Tyr
225 230 235 240

Ala Gly Leu Asn Val His Arg Gln
245

<210> 31

<211> 248

<212> PRT

<213> Escherichia coli

<400> 31

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

Thr Arg Arg Phe Ile Gln Gln Gly His Lys Val Ile Ala Thr Gly Arg
20 25 30

Arg Gln Glu Arg Leu Gln Glu Leu Lys Asp Glu Leu Gly Asp Asn Leu
35 40 45

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

Leu Ala Ser Leu Pro Ala Glu Trp Cys Asn Ile Asp Ile Leu Val Asn
65 70 75 80

Asn Ala Gly Leu Ala Leu Gly Met Glu Pro Ala His Lys Ala Ser Val
85 90 95

Glu Asp Trp Glu Thr Met Ile Asp Thr Asn Asn Lys Gly Leu Val Tyr
100 105 110

Met Thr Arg Ala Val Leu Pro Gly Met Val Glu Arg Asn His Gly His
115 120 125

Ile Ile Asn Ile Gly Ser Thr Ala Gly Ser Trp Pro Tyr Ala Gly Gly
130 135 140

Asn Val Tyr Gly Ala Thr Lys Ala Phe Val Arg Gln Phe Ser Leu Asn
145 150 155 160

Leu Arg Thr Asp Leu His Gly Thr Ala Val Arg Val Thr Asp Ile Glu
165 170 175

Pro Gly Leu Val Gly Gly Thr Glu Phe Ser Asn Val Arg Phe Lys Gly
180 185 190

Asp Asp Gly Lys Ala Glu Lys Thr Tyr Gln Asn Thr Val Ala Leu Thr
195 200 205

Pro Glu Asp Val Ser Glu Ala Val Trp Trp Val Ser Thr Leu Pro Ala
210 215 220

```

His Val Asn Ile Asn Thr Leu Glu Met Met Pro Val Thr Gln Ser Tyr
225                230                235                240

```

```

Ala Gly Leu Asn Val His Arg Gln
                245

```

<210> 32

<211> 248

<212> PRT

<213> Escherichia coli

<400> 32

```

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

```

```

Thr Arg Arg Phe Ile Gln Gln Gly His Lys Val Ile Ala Thr Asp Arg
          20           25           30

```

```

Arg Gln Glu Arg Leu Gln Glu Leu Lys Asp Glu Leu Gly Asp Asn Leu
          35           40           45

```

```

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

```

```

Leu Ala Ser Leu Pro Ala Glu Trp Cys Asn Ile Asp Ile Leu Val Asn
65           70           75           80

```

```

Asn Ala Gly Leu Ala Leu Gly Met Glu Pro Ala His Lys Ala Ser Val
          85           90           95

```

```

Glu Asp Trp Glu Thr Met Ile Asp Thr Asn Asn Lys Gly Leu Val Tyr
          100          105          110

```

```

Met Thr Arg Ala Val Leu Pro Gly Met Val Glu Arg Asn His Gly His
          115          120          125

```

```

Ile Ile Asn Ile Gly Ser Thr Ala Gly Ser Trp Pro Tyr Ala Gly Gly
          130          135          140

```

```

Asn Val Tyr Gly Ala Thr Lys Ala Phe Val Arg Gln Phe Ser Leu Asn
145          150          155          160

```

```

Leu Arg Thr Asp Leu His Gly Thr Ala Val Arg Val Thr Asp Ile Glu
          165          170          175

```

```

Pro Gly Leu Val Gly Gly Thr Glu Phe Ser Asn Val Arg Phe Lys Gly
          180          185          190

```

```

Asp Asp Gly Lys Ala Glu Lys Thr Tyr Gln Asn Thr Val Ala Leu Thr
          195          200          205

```

```

Pro Glu Asp Val Ser Glu Ala Val Trp Trp Val Ser Thr Leu Pro Ala
          210          215          220

```

```

His Val Asn Ile Asn Thr Leu Glu Met Met Pro Val Thr Gln Ser Tyr
225                230                235                240

```

```

Ala Gly Leu Asn Val His Arg Gln
                245

```

<210> 33

<211> 248

<212> PRT

<213> Escherichia coli

<400> 33

```

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

Thr Arg Arg Phe Ile Gln Gln Gly His Lys Val Ile Ala Thr Asp Leu
 20             25             30

Arg Gln Glu Arg Leu Gln Glu Leu Lys Asp Glu Leu Gly Asp Asn Leu
 35             40             45

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

Leu Ala Ser Leu Pro Ala Glu Trp Cys Asn Ile Asp Ile Leu Val Asn
 65             70             75             80

Asn Ala Gly Leu Ala Leu Gly Met Glu Pro Ala His Lys Ala Ser Val
 85             90             95

Glu Asp Trp Glu Thr Met Ile Asp Thr Asn Asn Lys Gly Leu Val Tyr
100            105            110

Met Thr Arg Ala Val Leu Pro Gly Met Val Glu Arg Asn His Gly His
115            120            125

Ile Ile Asn Ile Gly Ser Thr Ala Gly Ser Trp Pro Tyr Ala Gly Gly
130            135            140

Asn Val Tyr Gly Ala Thr Lys Ala Phe Val Arg Gln Phe Ser Leu Asn
145            150            155            160

Leu Arg Thr Asp Leu His Gly Thr Ala Val Arg Val Thr Asp Ile Glu
165            170            175

Pro Gly Leu Val Gly Gly Thr Glu Phe Ser Asn Val Arg Phe Lys Gly
180            185            190

Asp Asp Gly Lys Ala Glu Lys Thr Tyr Gln Asn Thr Val Ala Leu Thr
195            200            205

Pro Glu Asp Val Ser Glu Ala Val Trp Trp Val Ser Thr Leu Pro Ala
210            215            220

His Val Asn Ile Asn Thr Leu Glu Met Met Pro Val Thr Gln Ser Tyr
225            230            235            240

Ala Gly Leu Asn Val His Arg Gln
245

```

<210> 34

<211> 43

<212> DNA

<213> Escherichia coli

<400> 34

tcgccactga tctgaacccg gaagcgttgc aggagttaaa aga

43

<210> 35

<211> 43
<212> DNA
<213> Escherichia coli

<400> 35
tctttaact cctgcaacgc ttccgggttc agatcagtg cga 43

<210> 36
<211> 43
<212> DNA
<213> Escherichia coli

<400> 36
ccactgatct gaacccggcc cggttgcagg agttaaaga cga 43

<210> 37
<211> 43
<212> DNA
<213> Escherichia coli

<400> 37
tcgtcttta actcctgcaa ccgggcccggg ttcagatcag tgg 43

<210> 38
<211> 43
<212> DNA
<213> Escherichia coli

<400> 38
ggcataaagt tatcgccact ggctgaacc cggccgcgtt gca 43

<210> 39
<211> 43
<212> DNA
<213> Escherichia coli

<400> 39
tgcaacgcgg ccgggttcag gccagtggcg ataatcttat gcc 43

<210> 40
<211> 44
<212> DNA
<213> Escherichia coli

<400> 40
tcgttttagt aactggagca acggcaggtt ttggtgaatg catt 44

<210> 41
<211> 44
<212> DNA
<213> Escherichia coli

<400> 41
aatgattca ccaaacctg ccgttgctcc agttactaaa acga 44

<210> 42
<211> 43
<212> DNA
<213> Escherichia coli

<400> 42
ataaagttat cgccactgat cgtaaccggg ccgcgttgca gga 43

<210> 43
<211> 43
<212> DNA
<213> Escherichia coli

<400> 43
tctgcaacg cggccgggtt acgatcagtg gcgataactt tat 43

<210> 44
<211> 43
<212> DNA
<213> Escherichia coli

<400> 44
aactctgca acgcggccgg gcgcagatca gtggcgataa ctt 43

<210> 45
<211> 43
<212> DNA
<213> Escherichia coli

<400> 45
aagttatcgc cactgatctg cgcccggccg cgttcagga gtt 43

<210> 46
<211> 43
<212> DNA
<213> Escherichia coli

<400> 46
ttatgccac tgatctgaac caggccgctg tgcaggagtt aaa 43

<210> 47
<211> 43
<212> DNA
<213> Escherichia coli

<400> 47
ttaaactct gcaacgcggc ctggttcaga tcagtggcga taa 43

<210> 48
<211> 53
<212> DNA
<213> Escherichia coli

<400> 48

taaagttatc gccactgatc tgcgcccgga agcgttcag gagttaaag acg 53

<210> 49

<211> 53

<212> DNA

<213> Escherichia coli

<400> 49

cgtcttttaa ctctgcaac gcttcgggc gcagatcagt ggcgataact tta 53

<210> 50

<211> 53

<212> DNA

<213> Escherichia coli

<400> 50

agttatgcc actgatctgc gccaggccgc gttgcaggag taaaagacg aac 53

<210> 51

<211> 53

<212> DNA

<213> Escherichia coli

<400> 51

gttcgtcttt taactcctgc aacgcggcct ggcgcagatc agtggcgata act 53

<210> 52

<211> 59

<212> DNA

<213> Escherichia coli

<400> 52

gcataaagtt atcgccactg atctgcgcca ggaagcgttg caggagttaa aagacgaac 59

<210> 53

<211> 59

<212> DNA

<213> Escherichia coli

<400> 53

gttcgtcttt taactcctgc aacgcttct ggcgcagatc agtggcgata acttatgc 59

<210> 54

<211> 53

<212> DNA

<213> Escherichia coli

<400> 54

gcataaagtt atcgccactg atctgaacca ggaagcgttg caggagttaa aag 53

<210> 55

<211> 53

<212> DNA

<213> Escherichia coli

<400> 55
ctttaactc ctgcaacgt tctggttca gatcagtggc gataacttta tgc 53

<210> 56
<211> 40
<212> DNA
<213> Escherichia coli

<400> 56
gtcatcgtag tctagataaa atgatcgttt tggtcaccgg 40

<210> 57
<211> 28
<212> DNA
<213> Escherichia coli

<400> 57
gtgtccatt aattaattat tgtctgtg 28

<210> 58
<211> 30
<212> DNA
<213> Escherichia coli

<400> 58
gcggaattca tgttggtaa tatttccaa 30

<210> 59
<211> 29
<212> DNA
<213> Escherichia coli

<400> 59
gatcccgggc tatttatcta atgacctc 29

<210> 60
<211> 39
<212> DNA
<213> Escherichia coli

<400> 60
gttttagtaa ctggagcagg cgcaggttt ggtgaatgc 39

<210> 61
<211> 39
<212> DNA
<213> Escherichia coli

<400> 61
gcattacca aaacctgcgc ctgctccagt tactaaaac 39

<210> 62
<211> 39
<212> DNA

<213> Escherichia coli

<400> 62

cataaagtta tcgccactga tcgtgccag gaacggttg 39

<210> 63

<211> 39

<212> DNA

<213> Escherichia coli

<400> 63

caaccgttcc tggcgacgat cagtggcgat aactttatg 39

<210> 64

<211> 40

<212> DNA

<213> Escherichia coli

<400> 64

taaagttatc gccactgatc gtcgccagga agcgttgag 40

<210> 65

<211> 40

<212> DNA

<213> Escherichia coli

<400> 65

ctgcaacgct tctggcgac gatcagtgga gataactta 40

<210> 66

<211> 39

<212> DNA

<213> Escherichia coli

<400> 66

actgatctgc gccaggaacg gttgcaggag taaaagac 39

<210> 67

<211> 39

<212> DNA

<213> Escherichia coli

<400> 67

gtctttaac tctgcaacc gttctggcg cagatcagt 39

<210> 68

<211> 39

<212> DNA

<213> Escherichia coli

<400> 68

atcctaatta caggtgcggg tactggtatc ggataccat 39

<210> 69

<211> 39

<212> DNA

<213> Escherichia coli

<400> 69

atggtatccg ataccagtac ccgcacctgt aattaggat 39

<210> 70

<211> 42

<212> DNA

<213> Escherichia coli

<400> 70

ttgaagttgg ttttggtga ttaagaaag gagaagctgg ag 42

<210> 71

<211> 42

<212> DNA

<213> Escherichia coli

<400> 71

ctccagcttc tccttctta aatcagccaa aaccaacttc aa 42

<210> 72

<211> 31

<212> DNA

<213> Escherichia coli

<400> 72

ttgcaggcaa gaacatccta attacagtg c 31

<210> 73

<211> 31

<212> DNA

<213> Escherichia coli

<400> 73

gcacctgtaa ttaggatgtt ctgcctgca a 31

<210> 74

<211> 33

<212> DNA

<213> Escherichia coli

<400> 74

gtagctagct aaaatgttg gtaatatttc cca 33

<210> 75

<211> 31

<212> DNA

<213> Escherichia coli

<400> 75

tgcttaatta actatttadc taatgatcct c 31

<210> 76

<211> 33

<212> DNA

<213> Escherichia coli

<400> 76

ttggtcaccg gtgcaggtgc aggtttcggc gaa 33

<210> 77

<211> 33

<212> DNA

<213> Escherichia coli

<400> 77

ttgccgaaa cctgcacctg caccggtgac caa 33

<210> 78

<211> 40

<212> DNA

<213> Escherichia coli

<400> 78

acaaggttat cgctaccgac ttgagacaag agagattgca 40

<210> 79

<211> 40

<212> DNA

<213> Escherichia coli

<400> 79

tgcaatctct cttgtctcaa gtcggtagcg ataaccttgt 40

<210> 80

<211> 270

<212> PRT

<213> Issatchenkia orientalis

<400> 80

Met	Phe	Gly	Asn	Ile	Ser	Gln	Arg	Leu	Ala	Gly	Lys	Asn	Ile	Leu	Ile
1			5					10					15		

Thr	Gly	Ala	Ser	Thr	Gly	Ile	Gly	Tyr	His	Thr	Ala	Lys	Tyr	Phe	Ala
		20				25						30			

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

Glu	Lys	Leu	Glu	Ala	Leu	Lys	Ala	Asp	Leu	Leu	Ala	Lys	Tyr	Pro	Ser
	50				55						60				

Ile	Lys	Val	His	Ile	Glu	Ser	Leu	Asp	Val	Ser	Lys	Thr	Glu	Thr	Ile
65				70					75					80	

Ala	Pro	Phe	Leu	Lys	Gly	Leu	Pro	Glu	Glu	Phe	Ser	Ile	Val	Asp	Val
			85					90						95	

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

Val Asp Pro Lys Asp Val Asp Glu Met Phe Gln Thr Asn Val Leu Gly
115 120 125

Met Ile Gln Leu Thr Gln Leu Val Val Gln Gln Met Lys Glu Arg Asn
130 135 140

Ser Gly Asp Ile Val Gln Leu Gly Ser Val Ala Gly Arg Asn Pro Tyr
145 150 155 160

Pro Gly Gly Gly Ile Tyr Cys Ala Ser Lys Ala Ala Leu Arg Ser Phe
165 170 175

Thr His Val Leu Arg Glu Glu Leu Ile Asn Thr Lys Ile Arg Val Ile
180 185 190

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

Phe Lys Gly Asp Lys Ser Lys Ala Glu Lys Val Tyr Glu Gly Thr Glu
210 215 220

Pro Leu Tyr Gly Thr Asp Ile Ala Glu Leu Ile Leu Phe Ala Val Ser
225 230 235 240

Arg Pro Gln Asn Thr Val Ile Ala Glu Thr Leu Val Phe Ala Ser Asn
245 250 255

Gln Ala Ser Ala Tyr His Ile Phe Arg Gly Ser Leu Asp Lys
260 265 270

<210> 81

<211> 270

<212> PRT

<213> Issatchenkia orientalis

<400> 81

Met Phe Gly Asn Ile Ser Gln Arg Leu Ala Gly Lys Asn Ile Leu Ile
1 5 10 15

Thr Gly Ala Gly Thr Gly Ile Gly Tyr His Thr Ala Lys Tyr Phe Ala
20 25 30

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

Glu Lys Leu Glu Ala Leu Lys Ala Asp Leu Leu Ala Lys Tyr Pro Ser
50 55 60

Ile Lys Val His Ile Glu Ser Leu Asp Val Ser Lys Thr Glu Thr Ile
65 70 75 80

Ala Pro Phe Leu Lys Gly Leu Pro Glu Glu Phe Ser Ile Val Asp Val
85 90 95

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

Val Asp Pro Lys Asp Val Asp Glu Met Phe Gln Thr Asn Val Leu Gly
115 120 125

Met Ile Gln Leu Thr Gln Leu Val Val Gln Gln Met Lys Glu Arg Asn
130 135 140

Ser Gly Asp Ile Val Gln Leu Gly Ser Val Ala Gly Arg Asn Pro Tyr
145 150 155 160

Pro Gly Gly Gly Ile Tyr Cys Ala Ser Lys Ala Ala Leu Arg Ser Phe
165 170 175

Thr His Val Leu Arg Glu Glu Leu Ile Asn Thr Lys Ile Arg Val Ile
180 185 190

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

Phe Lys Gly Asp Lys Ser Lys Ala Glu Lys Val Tyr Glu Gly Thr Glu
210 215 220

Pro Leu Tyr Gly Thr Asp Ile Ala Glu Leu Ile Leu Phe Ala Val Ser
225 230 235 240

Arg Pro Gln Asn Thr Val Ile Ala Glu Thr Leu Val Phe Ala Ser Asn
245 250 255

Gln Ala Ser Ala Tyr His Ile Phe Arg Gly Ser Leu Asp Lys
260 265 270

<210> 82

<211> 295

<212> PRT

<213> Issatchenkia orientalis

<400> 82

Met Arg Ile Ala Phe Ile Gly Leu Gly Asn Met Gly Ala Pro Met Ala
1 5 10 15

Arg Asn Leu Ile Lys Ala Gly His Gln Leu Asn Leu Phe Asp Leu Asn
20 25 30

Lys Ala Val Leu Ala Glu Leu Ala Glu Leu Gly Gly Gln Ile Ser Pro
35 40 45

Ser Pro Lys Asp Ala Ala Ala Asn Ser Glu Leu Val Ile Thr Met Leu
50 55 60

Pro Ala Ala Ala His Val Arg Ser Val Tyr Leu Asn Glu Asp Gly Val
65 70 75 80

Leu Ala Gly Ile Arg Pro Gly Thr Pro Thr Val Asp Cys Ser Thr Ile
85 90 95

Asp Pro Gln Thr Ala Arg Asp Val Ser Lys Ala Ala Ala Ala Lys Gly
100 105 110

Val Asp Met Gly Asp Ala Pro Val Ser Gly Gly Thr Gly Gly Ala Ala
115 120 125

Ala Gly Thr Leu Thr Phe Met Val Gly Ala Ser Thr Glu Leu Phe Ala
130 135 140

Ser Leu Lys Pro Val Leu Glu Gln Met Gly Arg Asn Ile Val His Cys
145 150 155 160

Gly Glu Val Gly Thr Gly Gln Ile Ala Lys Ile Cys Asn Asn Leu Leu
165 170 175

```

Leu Gly Ile Ser Met Ile Gly Val Ser Glu Ala Met Ala Leu Gly Asn
    180                      185                      190

Ala Leu Gly Ile Asp Thr Lys Val Leu Ala Gly Ile Ile Asn Ser Ser
    195                      200                      205

Thr Gly Arg Cys Trp Ser Ser Asp Thr Tyr Asn Pro Trp Pro Gly Ile
    210                      215                      220

Ile Glu Thr Ala Pro Ala Ser Arg Gly Tyr Thr Gly Gly Phe Gly Ala
    225                      230                      235                      240

Glu Leu Met Leu Lys Asp Leu Gly Leu Ala Thr Glu Ala Ala Arg Gln
    245                      250                      255

Ala His Gln Pro Val Ile Leu Gly Ala Val Ala Gln Gln Leu Tyr Gln
    260                      265                      270

Ala Met Ser Leu Arg Gly Glu Gly Gly Lys Asp Phe Ser Ala Ile Val
    275                      280                      285

Glu Gly Tyr Arg Lys Lys Asp
    290                      295

```

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- [WO2008027742A1 \[0003\]](#)
- [WO2012074818A \[0045\] \[0060\] \[0131\] \[0131\] \[0152\] \[0152\] \[0167\] \[0174\] \[0183\] \[0184\] \[0184\]](#)
- [EP238023A \[0059\]](#)
- [WO9600787A \[0059\] \[0081\]](#)
- [WO0242418A \[0060\]](#)
- [WO2008027742A \[0060\]](#)
- [US20040171154A \[0070\]](#)
- [WO9517413A \[0073\]](#)
- [WO9522625A \[0073\]](#)
- [US5223409A \[0073\]](#)
- [WO9206204A \[0073\]](#)
- [WO9943835A \[0080\]](#)
- [WO0056900A \[0081\] \[0081\] \[0081\]](#)
- [WO9425612A \[0088\]](#)
- [WO9533836A \[0099\]](#)
- [WO0024883A \[0112\] \[0112\]](#)
- [US6154136B \[0186\]](#)

Non-patent literature cited in the description

- **NEEDLEMANWUNSCH**J. Mol. Biol., 1970, vol. 48, 443-453 [0025] [0030]
- **RICE et al.**EMBOSS: The European Molecular Biology Open Software SuiteTrends Genet., 2000, vol. 16, 276-277 [0025] [0030]
- **RICE et al.**EMBOSS: The European Molecular Biology Open Software Suite, 2000, [0026]
- **EDGAR**Nucleic Acids Research, 2004, vol. 32, 1792-1797 [0031]
- **KATOHKUMAN**Nucleic Acids Research, 2002, vol. 30, 3059-3066 [0031]
- **KATOH et al.**Nucleic Acids Research, 2005, vol. 33, 511-518 [0031]
- **KATOHTOH**Bioinformatics, 2007, vol. 23, 372-374 [0031]
- **KATOH et al.**Methods in Molecular Biology, 2009, vol. 537, 39-64 [0031]
- **KATOHTOH**Bioinformatics, 2010, vol. 26, 1899-1900 [0031]
- **THOMPSON et al.**Nucleic Acids Research, 1994, vol. 22, 4673-4680 [0031]
- **LINDAHLELOFSSON**J. Mol. Biol., 2000, vol. 295, 613-615 [0032]
- **ATSCHUL et al.**Nucleic Acids Res., 1997, vol. 25, 3389-3402 [0032]
- **JONES**J. Mol. Biol., 1999, vol. 287, 797-815 [0032]
- **MCGUFFINJONES**Bioinformatics, 2003, vol. 19, 874-881 [0032]
- **GOUGH et al.**J. Mol. Biol., 2000, vol. 313, 903-919 [0032]
- **HOLMSANDER**Proteins, 1998, vol. 33, 88-96 [0033]
- **SHINDYALOVBOURNE**Protein Engineering, 1998, vol. 11, 739-747 [0033]
- **HOLMPARK**Bioinformatics, 2000, vol. 16, 566-567 [0033]
- **CHANGCOHEN**Mol. Gen. Genet., 1979, vol. 168, 111-115 [0050]
- **YOUNGSPIZEN**J. Bacteriol., 1961, vol. 81, 823-829 [0050]
- **DUBNAUDAVIDOFF-ABELSON**J. Mol. Biol., 1971, vol. 56, 209-221 [0050]
- **SHIGEKAWADOWER**Biotechniques, 1988, vol. 6, 742-751 [0050]
- **KOEHLERTHORNE**J. Bacteriol., 1987, vol. 169, 5271-5278 [0050]
- **HANAHAN**J. Mol. Biol., 1983, vol. 166, 557-580 [0050]
- **DOWER et al.**Nucleic Acids Res., 1988, vol. 16, 6127-6145 [0050]
- **GONG et al.**Folia Microbiol. (Praha), 2004, vol. 49, 399-405 [0050]
- **MAZODIER et al.**J. Bacteriol., 1989, vol. 171, 3583-3585 [0050]
- **BURKE et al.**Proc. Natl. Acad. Sci. USA, 2001, vol. 98, 6289-6294 [0050]
- **CHOI et al.**J. Microbiol. Methods, 2006, vol. 64, 391-397 [0050]
- **PINEDOSMETS**Appl. Environ. Microbiol., 2005, vol. 71, 51-57 [0050]
- **PERRYKURAMITSU**Infect. Immun., 1981, vol. 32, 1295-1297 [0050]
- **CATTJOLICK**Microbios, 1991, vol. 68, 189-207 [0050]
- **BUCKLEY et al.**Appl. Environ. Microbiol., 1999, vol. 65, 3800-3804 [0050]
- **CLEWELL**Microbiol. Rev., 1981, vol. 45, 409-436 [0050]
- **HAWKSWORTH et al.**Ainsworth and Bisby's Dictionary of The FungiCAB International, University Press19950000 [0052]
- **Biology and Activities of Yeast**Soc. App. Bacteriol. Symposium Series No. 919800000 [0053]
- **YELTON et al.**Proc. Natl. Acad. Sci. USA, 1984, vol. 81, 1470-1474 [0059]
- **CHRISTENSEN et al.**Bio/Technology, 1988, vol. 6, 1419-1422 [0059]
- **MALARDIER et al.**Gene, 1989, vol. 78, 147-156 [0059]
- **Guide to Yeast Genetics and Molecular Biology**BECKERGUARENTEMethods in EnzymologyAcademic Press, Inc.vol. 194, 182-187 [0059]
- **ITO et al.**J. Bacteriol., 1983, vol. 153, 163- [0059]
- **HINNEN et al.**Proc. Natl. Acad. Sci. USA, 1978, vol. 75, 1920- [0059]

- **NOVICK** Proc Natl Acad Sci USA, 1950, vol. 36, 708-719 [0062]
- **HARDER** J Appl Bacteriol, 1977, vol. 43, 1-24 [0062]
- **SAMBROOK et al.** Molecular Cloning: A Laboratory Manual Cold Spring Harbor Laboratory 20010000 [0065]
- **AUSUBEL et al.** Current Protocols in Molecular Biology John Wiley and Sons 19990000 [0065]
- **HANAI et al.** Appl. Environ. Microbiol., 2007, vol. 73, 7814-7818 [0065]
- **SCHERER DAVIS** Proc. Natl. Acad. Sci. USA, 1979, vol. 76, 4949-4955 [0069]
- **BARTON et al.** Nucleic Acids Res., 1990, vol. 18, 7349-4966 [0069]
- **STORICI et al.** Nature Biotechnol., 2001, vol. 19, 773-776 [0070]
- **KREN et al.** Nat. Med., 1998, vol. 4, 285-290 [0070]
- **CALISSANOMACINO** Fungal Genet. Newslett., 1996, vol. 43, 15-16 [0070]
- **TIAN et al.** Nature, 2004, vol. 432, 1050-1054 [0072]
- **REIDHAAR-OLSON SAUER** Science, 1988, vol. 241, 53-57 [0073]
- **BOWIE SAUER** Proc. Natl. Acad. Sci. USA, 1989, vol. 86, 2152-2156 [0073]
- **LOWMAN et al.** Biochemistry, 1991, vol. 30, 10832-10837 [0073]
- **DERBYSHIRE et al.** Gene, 1986, vol. 46, 145- [0073]
- **NER et al.** DNA, 1988, vol. 7, 127- [0073]
- **NESS et al.** Nature Biotechnology, 1999, vol. 17, 893-896 [0074]
- **AGAISSELERECLUS** Molecular Microbiology, 1994, vol. 13, 97-107 [0080]
- **EGON et al.** Gene, 1988, vol. 69, 301-315 [0080]
- **VILLA-KAMAROFF et al.** Proc. Natl. Acad. Sci. USA, 1978, vol. 75, 3727-3731 [0080]
- **DEBOER et al.** Proc. Natl. Acad. Sci. USA, 1983, vol. 80, 21-25 [0080]
- **GILBERT et al.** Useful proteins from recombinant bacteria Scientific American, 1980, vol. 242, 74-94 [0080]
- **ROMANOS et al.** Yeast, 1992, vol. 8, 423-488 [0082]
- **HUE et al.** Journal of Bacteriology, 1995, vol. 177, 3465-3471 [0088]
- **GUOSHHERMAN** Mol. Cellular Biol., 1995, vol. 15, 5983-5990 [0094]
- **SIMONENPALVA** Microbiological Reviews, 1993, vol. 57, 109-137 [0096]
- **GEMS et al.** Gene, 1991, vol. 98, 61-67 [0112]
- **CULLEN et al.** Nucleic Acids Res., 1987, vol. 15, 9163-9175 [0112]
- **LIN et al.** Biotechnol. Bioeng., 2005, vol. 90, 775-779 [0127]
- **AMANN, E. et al.** Tightly regulated tac promoter vectors useful for the expression of unfused and fused proteins in Escherichia coli Gene, 1988, vol. 69, 2301-315 [0134]
- **SHEEN, J.** High-Efficiency Transformation by Electroporation Current Protocols in Molecular Biology, 1989, vol. 1, 84- [0146]
- **YAMADA JACOBY** Direct conversion of malonic semialdehyde to acetyl-coenzyme A J. Biol. Chem., 1960, vol. 235, 3589-594 [0150]

KRAV

1. Værtscelle, som omfatter en aktiv 3-hydroxypropionat-vej og et heterologt polynukleotid, der koder for en 3-hydroxypropionat-dehydrogenasevariant med 3-hydroxypropionat-dehydrogenaseaktivitet, hvor varianten omfatter eller består af SEQ ID NO: 10, 12, 20, 21, 22, 23, 24, 25, 26, 27, 29, 30, 32 eller 81.
2. Værtscelle ifølge krav 1, hvor varianten har forhøjet specificitet for NAD(H) sammenlignet med NADP(H) (f.eks. større end 2 gange, 5 gange, 10 gange, 20 gange, 50 gange, 100 gange, 200 gange, 500 gange eller 1000-gange specificitet for NAD(H) sammenlignet med NADP(H)).
3. Værtscelle ifølge krav 1 eller 2, hvor cellen er en gærcele (f.eks. en *Issatchenkia orientalis*-gærcele).
4. Værtscelle ifølge ethvert af kravene 1-3, hvor cellen omfatter: PEP-carboxylaseaktivitet eller pyruvatcarboxylaseaktivitet; aspartat-aminotransferaseaktivitet; aspartat-decarboxylaseaktivitet; og beta-alanin/alpha-ketoglutarat-aminotransferase (BAAT)-aktivitet.
5. Værtscelle ifølge ethvert af kravene 1-4, hvor værtscellen omfatter et eller flere heterologe polynukleotider valgt fra:
et heterologt polynukleotid, der koder for en PEP-carboxylase,
et heterologt polynukleotid, der koder for en pyruvatcarboxylase,
et heterologt polynukleotid, der koder for en aspartat-aminotransferase,
et heterologt polynukleotid, der koder for en aspartat-decarboxylase, og
et heterologt polynukleotid, der koder for en BAAT.
6. Fremgangsmåde til fremstilling af 3-hydroxypropionat, omfattende:
 - a) dyrkning af værtscellen ifølge ethvert af kravene 1-5 under betingelser, der er befordrende for fremstilling af 3-hydroxypropionat; og
 - b) indvinding af 3-hydroxypropionat.

DRAWINGS

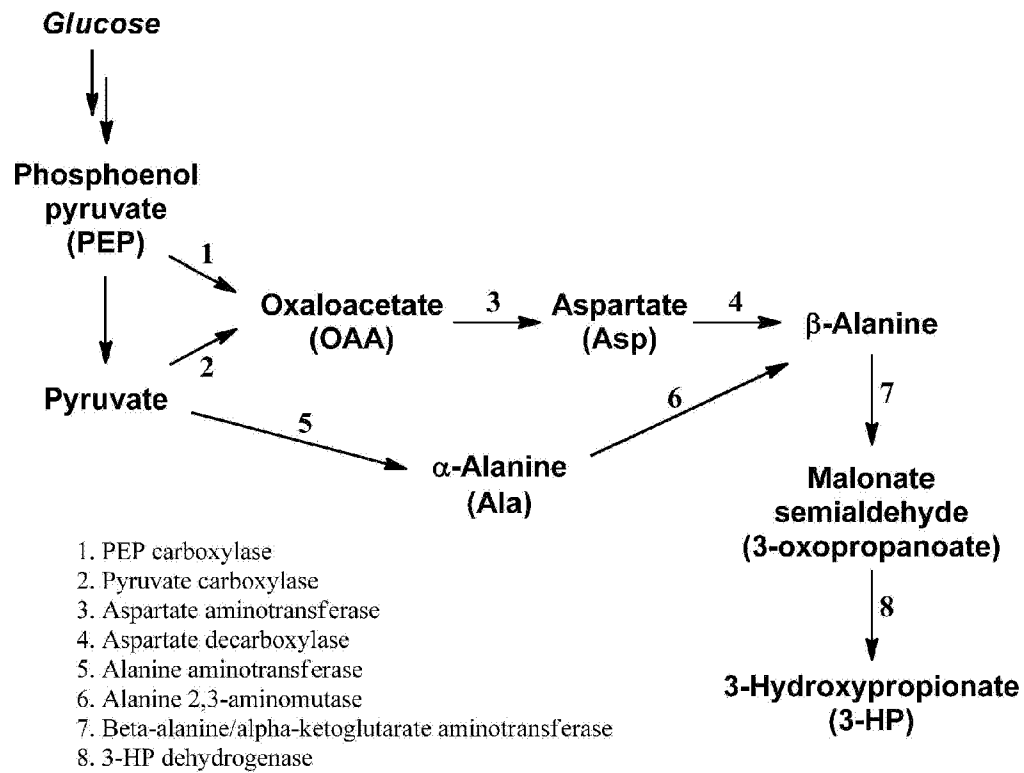


Fig. 1

<i>E. coli</i>	1	-----MIVLVIGATAGEGECITRRFIQQG---HKVIATGRRQERLQELKDEL
<i>I. orientalis</i>	1	MFG--NISQRLAGKNILITGASTGIGYHTAKYFAEAANGDLKLVLAARRKEKLEALKADL
<i>S. cerevisiae</i>	1	MSQGRKAAERLAKKTVLITGASAGIGKATALEYLEASNGDMKLIILAAARLEKLEELKKTII
<i>E. coli</i>	45	GDN-----LYIAQLDVRNRAAIEEMLASLPAEWCNIDILVNNAGLALGMEPAHKASVEDW
<i>I. orientalis</i>	59	LAKYPSIKVHIESLDSVSKTETIAPFLKGLPEEFSIVDVLVNNAGKALGLDPIGSVDPKDV
<i>S. cerevisiae</i>	61	DQEFFNAKVHVVAQLDITQAEKIKPFIENLPQEFKDDIDILVNNAGKALGSDRVGQIATEDI
<i>E. coli</i>	100	ETMIDTNNKGLVYMTRAVLPGMVERNHCHIIINIGTAGSWPYAGNVYGATKAFVRQPSL
<i>I. orientalis</i>	119	DEMFOINVLCMIQLTQLVVQOMKERNSGDIVQLGSAVGRNPYPGGGIYCASKAALRSFTH
<i>S. cerevisiae</i>	121	QDVFDTNVTALINITQAVLPFIQAKNSGDIVNLGSIAGRDAYPTGSIYCASKFAVGAFID
<i>E. coli</i>	160	NLRIDLHGTAVRVTDIEPGLVGGTEFSNVRFKGGDKAEKTYQNTVALTPEDVSEAVWV
<i>I. orientalis</i>	179	VLREELINTKIRVIEIEPGNVATEEFSLTFKFGDKSKAEKVYEGTEPLYGTDIAELILEA
<i>S. cerevisiae</i>	181	SLRKELINTKIRVILIAPG-LVETEEFLVRYRGNEEQAKNVYKDTTPLMADDDVADLIIVVA
<i>E. coli</i>	220	STLPAHVNINTLEMPVTSYAGLNVHRQ----
<i>I. orientalis</i>	239	VSRPQNTVIAETLVFASNQASA-YHIFRGS�DK
<i>S. cerevisiae</i>	240	TSRKQNTVIADTLIFPTNQASP-HHIFRG----

Fig. 2

E. coli (wt):
 mut1: MIVLVTGATAGFGGECITRRFIIQQG-- --HKVIATGRRQERLQELKDEL..
 mut2: MIVLVTGAG*GFGECITRRFIIQQG-- --HKVIATDLNPAALQELKDEL..
 mut3: MIVLVTGAGAGFGECITRRFIIQQG-- --HKVIATELNPAALQELKDEL..
 mut4: MIVLVTGAG*GFGECITRRFIIQQG-- --HKVIATELNPAALQELKDEL..
 mut5: MIVLVTGAGAGFGECITRRFIIQQG-- --HKVIATDLSADALQELKDEL..
 mut6: MIVLVTGAG*GFGECITRRFIIQQG-- --HKVIATDLSADALQELKDEL..
 mut7: MIVLVTGAGAGFGECITRRFIIQQG-- --HKVIATDLSADALQELKDEL..
 mut8: MIVLVTGAG*GFGECITRRFIIQQG-- --HKVIATELSADALQELKDEL..
 mut9: MIVLVTGATAGFGECITRRFIIQQG-- --HKVIATDLNPAALQELKDEL..
 mut10: MIVLVTGAGAGFGECITRRFIIQQG-- --HKVIATGILNPAALQELKDEL..
 mut11: MIVLVTGAGAGFGECITRRFIIQQG-- --HKVIATDRNPAALQELKDEL..
 mut12: MIVLVTGAGAGFGECITRRFIIQQG-- --HKVIATDLRPAALQELKDEL..
 mut13: MIVLVTGAGAGFGECITRRFIIQQG-- --HKVIATDLNQALQELKDEL..
 mut14: MIVLVTGAGAGFGECITRRFIIQQG-- --HKVIATDLNPAALQELKDEL..
 mut15: MIVLVTGAGAGFGECITRRFIIQQG-- --HKVIATDLNPAALQELKDEL..
 mut16: MIVLVTGAGAGFGECITRRFIIQQG-- --HKVIATDLRQALQELKDEL..
 mut17: MIVLVTGAGAGFGECITRRFIIQQG-- --HKVIATDLRQALQELKDEL..
 mut18: MIVLVTGAGAGFGECITRRFIIQQG-- --HKVIATDLRPAALQELKDEL..
 mut19: MIVLVTGAGAGFGECITRRFIIQQG-- --HKVIATDLNQALQELKDEL..
 mut20: MIVLVTGATAGFGECITRRFIIQQG-- --HKVIATDRRQERLQELKDEL..
 mut21: MIVLVTGAGAGFGECITRRFIIQQG-- --HKVIATDRRQERLQELKDEL..
 mut22: MIVLVTGAGAGFGECITRRFIIQQG-- --HKVIATDLRQERLQELKDEL..
 mut23: MIVLVTGAGAGFGECITRRFIIQQG-- --HKVIATGRRQERLQELKDEL..
 mut24: MIVLVTGAGAGFGECITRRFIIQQG-- --HKVIATDRRQERLQELKDEL..
 mut25: MIVLVTGATAGFGECITRRFIIQQG-- --HKVIATDLRQERLQELKDEL..
 I_o (wt): MFG--NISQRLAGKNILITGASTGIGYHTAKYFAEAANGDLKLVLAARRKEKLEALKADL..
 mut26: MFG--NISQRLAGKNILITGASTGIGYHTAKYFAEAANGDLKLVLDLRKEKLEALKADL..
 mut27: MFG--NISQRLAGKNILITGASTGIGYHTAKYFAEAANGDLKLVLDLRKEKLEALKADL..

Fig. 3

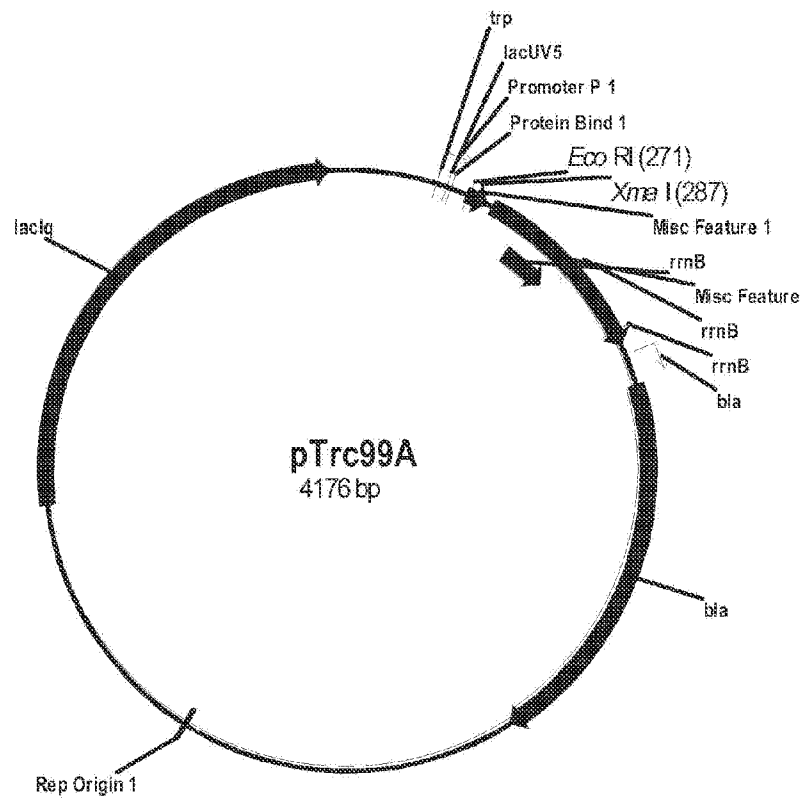


Fig. 4

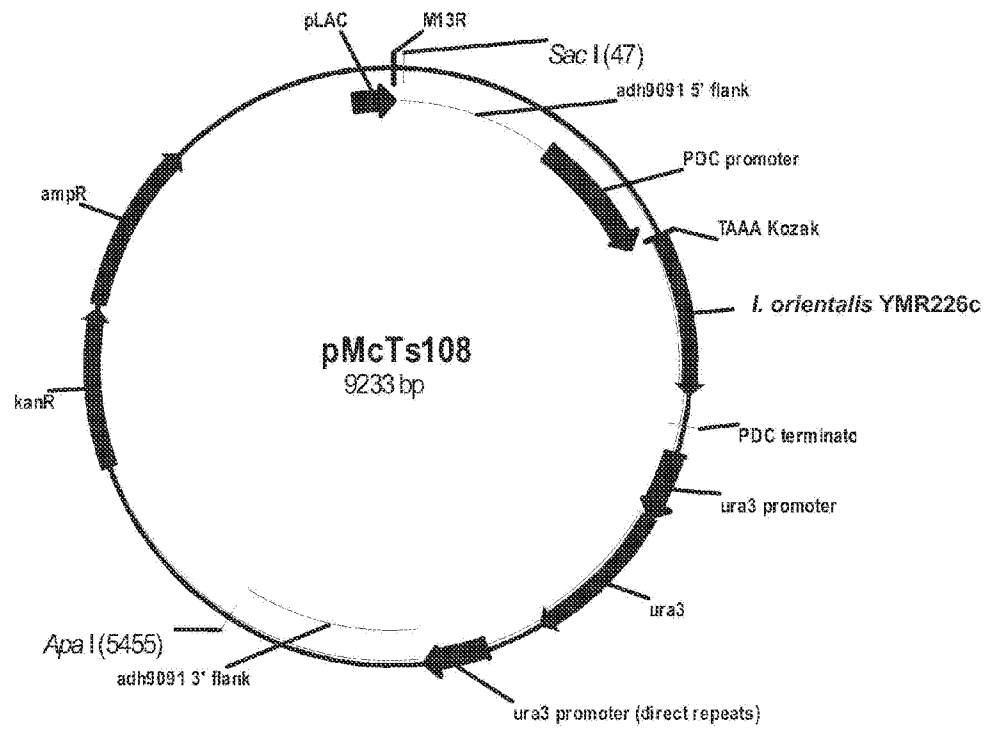


Fig. 5

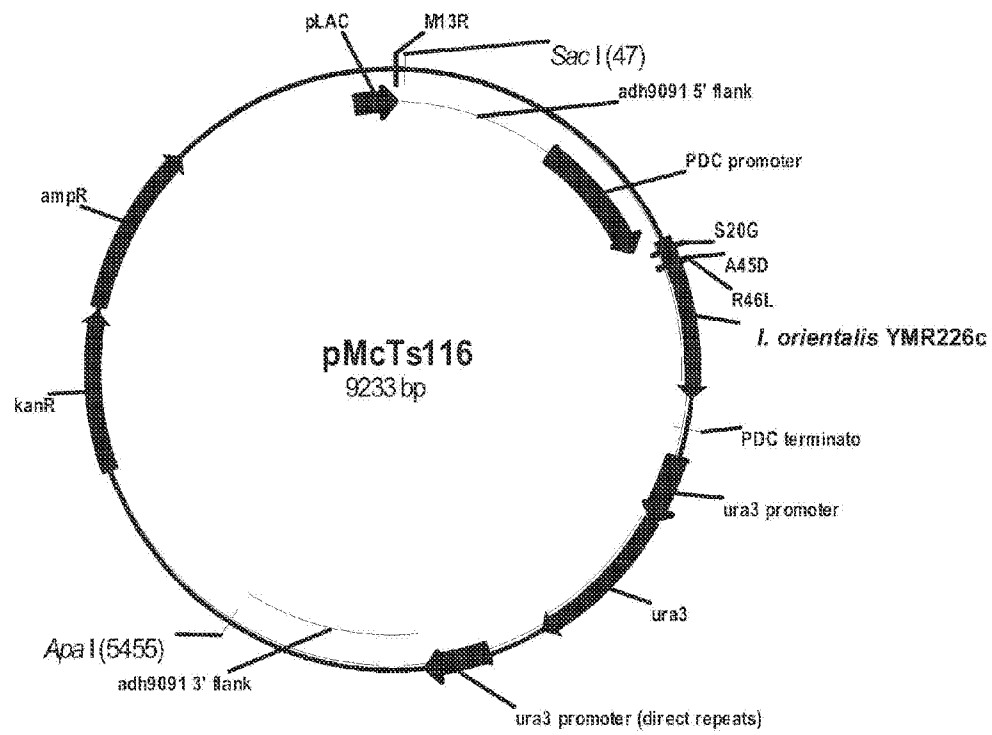


Fig. 6

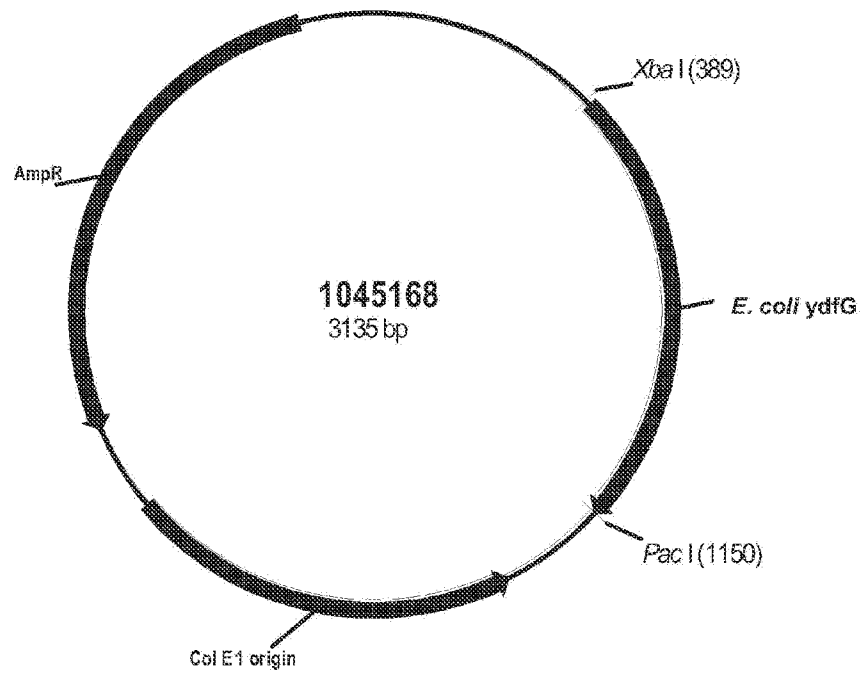


Fig. 7

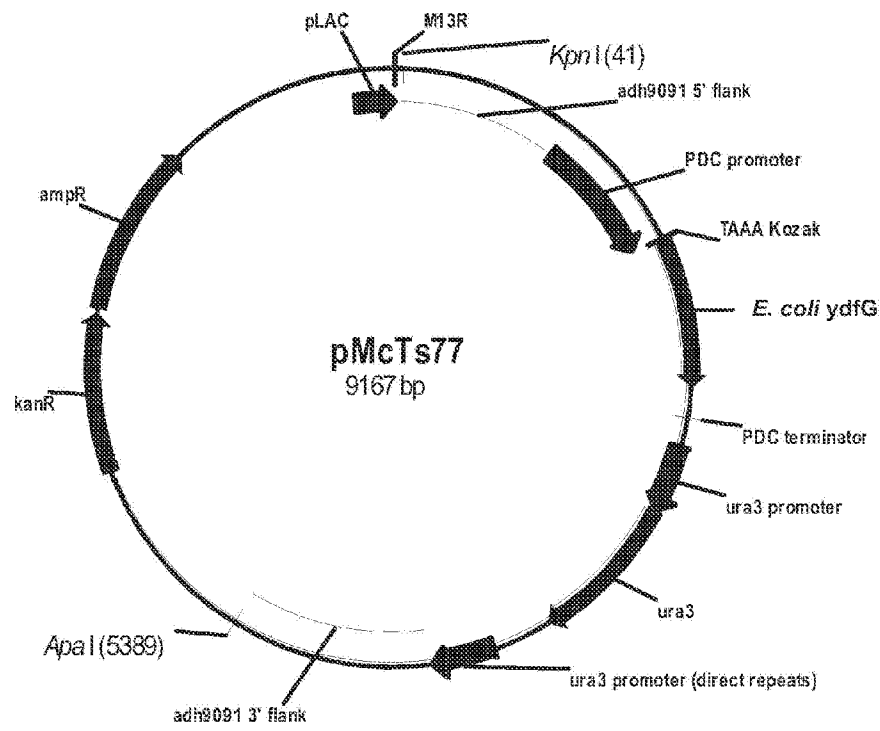


Fig. 8

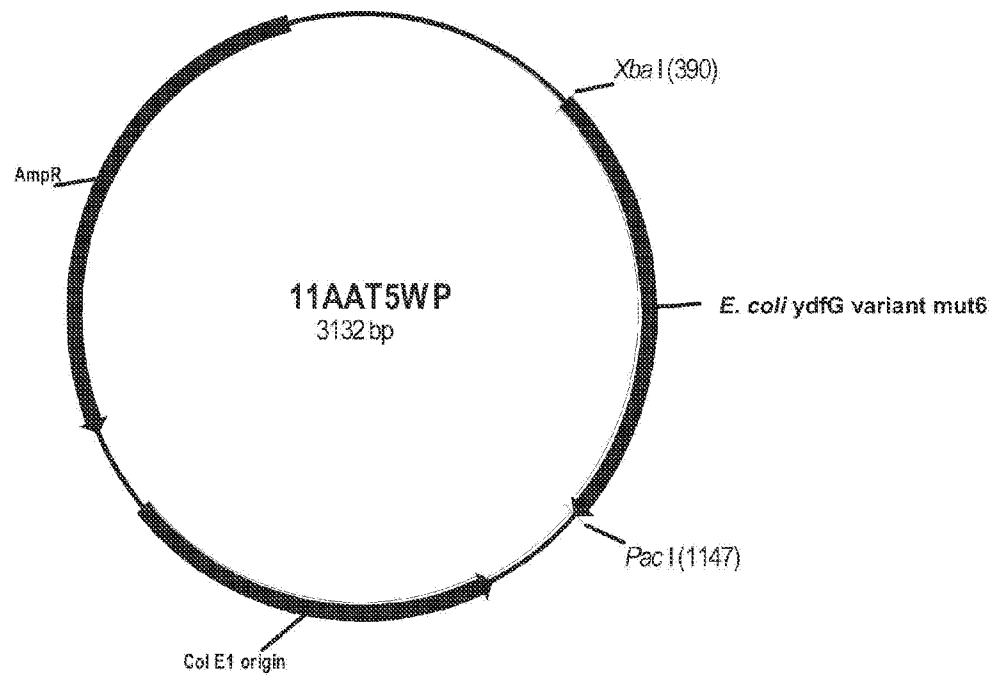


Fig. 9

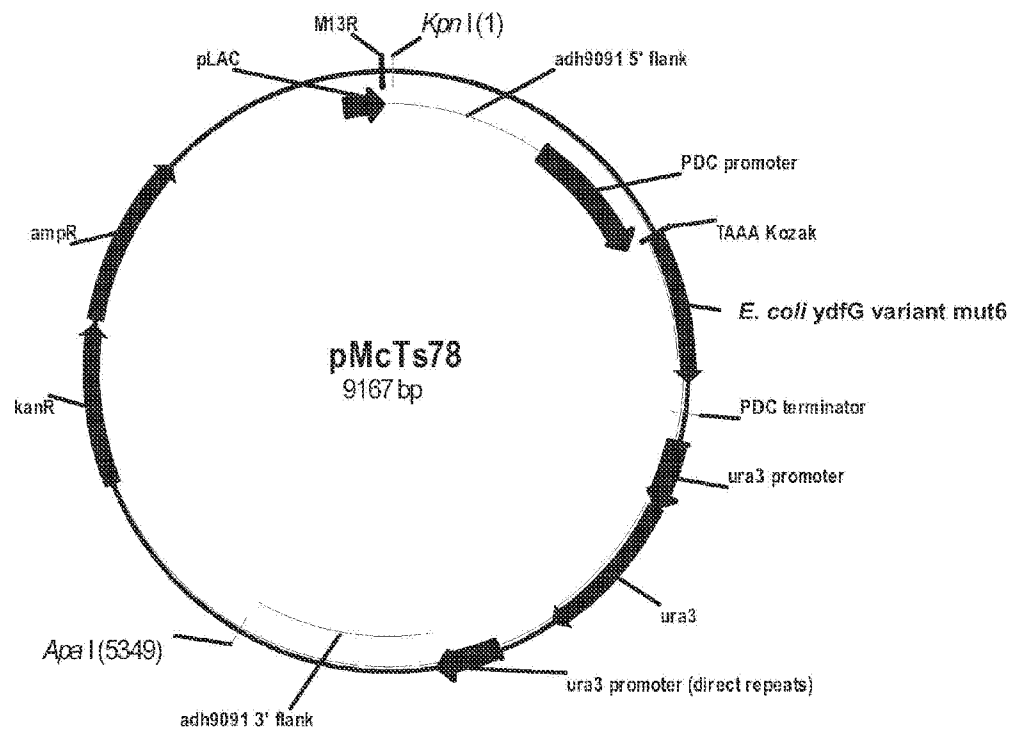


Fig. 10

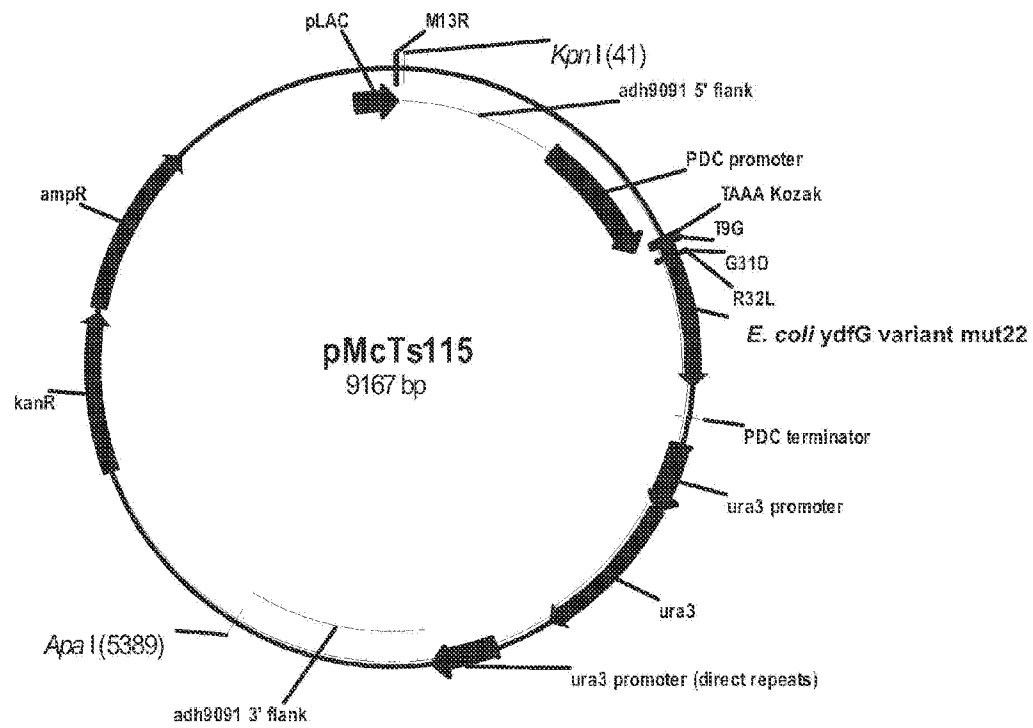


Fig. 11

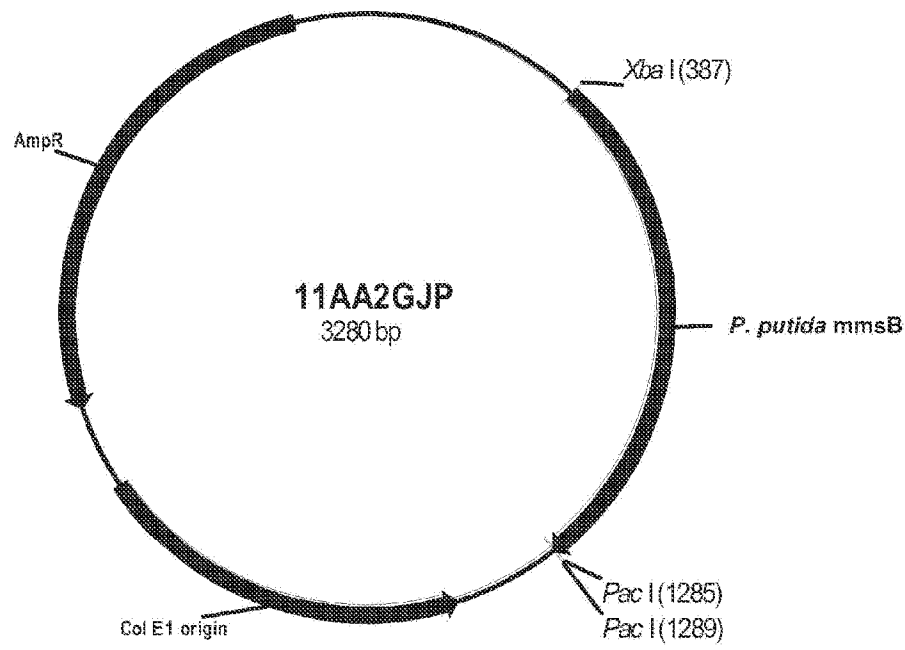


Fig. 12

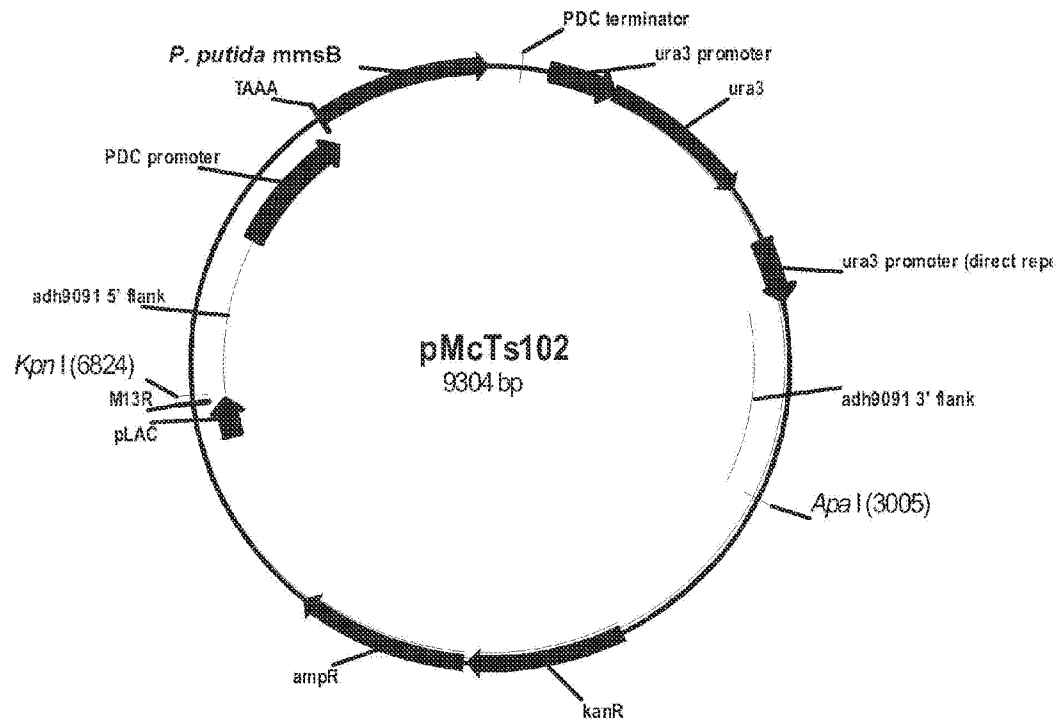


Fig. 13