(19) DANMARK

(10) DK/EP 3679059 T3

(12) Oversættelse af europæisk patentskrift

Patent- og Varemærkestyrelsen

(51) Int.Cl.: *C 07 K 14/435 (2006.01)* *C 12 N 9/90 (2006.01)* *C 12 P 19/24 (2006.01)*

(45) Oversættelsen bekendtgjort den: **2022-02-21**

(80) Dato for Den Europæiske Patentmyndigheds bekendtgørelse om meddelelse af patentet: **2021-11-17**

(86) Europæisk ansøgning nr.: **18762311.1**

(86) Europæisk indleveringsdag: **2018-09-04**

(87) Den europæiske ansøgnings publiceringsdag: **2020-07-15**

(86) International ansøgning nr.: **EP2018073747**

(87) Internationalt publikationsnr.: **WO2019043252**

(30) Prioritet: **2017-09-04 EP 17189270**

(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: **MetGen Oy, Rakentajantie 26, 20780 Kaarina, Finland**

(72) Opfinder: **BIRIKH, Klara, , 20780 KAARINA, Finland**
SUONPÄÄ, Anu Minna Maaret, , 20780 KAARINA, Finland

(74) Fuldmægtig i Danmark: **NORDIC PATENT SERVICE A/S, Bredgade 30, 1260 København K, Danmark**

(54) Benævnelse: **NYE FORBEDREDE GLUCOSEISOMERASER**

(56) Fremdragne publikationer:
WO-A1-2010/070549
WO-A1-2012/033926
WO-A1-2017/060195
Anonymous: "DICTH_1588 - Xylose isomerase, putative - Dictyoglomus thermophilum (strain ATCC 35947 / DSM 3960 / H-6-12) - DICTH_1588 gene & protein", , 25 November 2008 (2008-11-25), XP055413343, Retrieved from the Internet: URL:http://www.uniprot.org/uniprot/B5YAD2 [retrieved on 2017-10-06]

DESCRIPTION

Field of the invention

[0001] The invention is in the field of enzymology. More in particular, it provides a method for the isomerization of glucose into fructose wherein the glucose is derived from lignocellulosic material. It also provides an enzyme with an increased glucose isomerase activity in the presence of xylose. More in particular, the invention provides polypeptides encoding mutant glucose isomerase enzymes with improved glucose isomerase activity as compared to the corresponding wild type enzyme. The disclosed polypeptides are particularly suited for converting glucose to fructose in the presence of xylose.

Background of the invention

[0002] High-fructose corn syrup (HFCS) (also called glucose-fructose, isoglucose and glucose-fructose syrup) is a sweetener made from corn starch that has been processed by an enzyme called glucose isomerase (EC 5.3.1.5) to convert some of its glucose into fructose. HFCS was first marketed in the early 1970s by the Clinton Corn Processing Company, together with the Japanese research institute where the enzyme was discovered.

[0003] As a sweetener, HFCS is often compared to granulated sugar. Advantages of HFCS over granulated sugar include being easier to handle, and being less expensive in some countries. In the U.S., HFCS is among the sweeteners that mostly replaced sucrose (table sugar) in the food industry.

[0004] In a contemporary process, corn (maize) is milled to produce corn starch and an "acid-enzyme" process is used in which the corn starch solution is acidified to begin breaking up the existing carbohydrates, and then enzymes are added to further metabolize the starch and convert the resulting sugars to fructose.

[0005] Glucose isomerase is an enzyme which converts glucose to fructose in a reversible reaction with equilibrium around 1:1 ratio of glucose to fructose. The enzyme may be obtained from many different species of bacteria such as Streptomyces, Actinoplanes, Microbacterium and Bacillus, and the enzyme is or has been marketed by companies such as Enzyme Bio-systems, Genencor, Gist-Brocades, Solvay Enzyme Inc and Novo Nordisk.

[0006] Most successful commercial glucose isomerases are immobilized on a solid support and as a consequence are very stable with an extremely long half life. In a typical process, the immobilized isomerase is loaded in a column and substrate (feed stock) is passed through at a rate that produces an effluent containing 42% fructose. Prerequisite however, is that the feed stock is a refined hydrolysate containing 93-96% glucose. Efficient refining is required in order

to remove impurities that could cause inactivation of the glucose isomerase.

[0007] Glucose may also be obtained from lignocellulose material. The term "lignocellulose" refers to plant dry matter, so called lignocellulosic biomass. It is the most abundantly available carbon source on earth for the production of bio-fuels, mainly bio-ethanol and potentially bio-based materials such as polymer and plastics. It is composed of carbohydrate polymers (cellulose, hemicellulose), and an aromatic polymer (lignin). These carbohydrate polymers contain different sugar monomers (six-and five-carbon sugars) and they are tightly bound to lignin.

[0008] Use of the currently available glucose isomerases in the conversion of lignocellulose-derived glucose into fructose is hampered by the impurities that are present in lignocellulose-derived glucose. These impurities lead to a significant decrease in the stability of the enzyme and cause significant costs for feed stock purification.

[0009] Therefore, in order to avoid cumbersome and costly purification steps in the production of fructose from lignocellulose material, it is desirable to have a glucose isomerase that is resistant towards some or most, if not all impurities of lignocellulose-derived glucose, more specifically lignin and xylose.

[0010] In a co-pending application (European Patent Application EP 16175234.0) we identified a family of glucose isomerases derived from the genus of Diktyoglomus that were proven to be resistant against the decrease in stability caused by the presence of lignin.

[0011] Xylose is a largely preferred substrate for glucose isomerase, as compared to glucose. Thus, xylose competes with glucose for the enzyme and glucose conversion rate is thereby largely reduced. It is therefore desirable to have an enzyme with a glucose isomerase activity that is less or not at all inhibited by xylose.

Summary of the invention

[0012] The current invention discloses glucose isomerases that are resistant to inhibition by the presence of xylose in the reaction mixture and variants thereof. Thus the invention provides polypeptides encoding glucose isomerase enzymes with an increased or improved glucose isomerase activity.

[0013] The term "improved glucose isomerase activity" or "increased glucose isomerase activity" as used herein refers to an enzyme with a higher glucose isomerase activity as compared to a control enzyme. In other words, this means that the same amount of enzyme (expressed as mass of protein) is able to convert more glucose to fructose per minute as compared to a control enzyme.

[0014] More in particular, the invention provides a polypeptide with glucose isomerase activity

comprising an amino acid sequence that is at least 90% identical to the amino acid sequence according to SEQ ID NO: 1 or SEQ ID NO: 2, wherein the polypeptide comprises a tiny amino acid residue at an amino acid position corresponding to position 104 in SEQ ID NO: 1 or SEQ ID NO: 2.

[0015] This enzyme exhibits an improved glucose isomerase activity as compared to the control enzyme; i.e. wherein the control enzyme is the glucose isomerase according to SEQ ID NO: 1 or SEQ ID NO: 2 wherein the amino acid corresponding to position 104 is not a tiny amino acid.

[0016] In a preferred embodiment, the term "tiny amino acid" as used herein refers to amino acids selected from the group consisting of glycine, alanine, serine and cysteine.

[0017] The invention also relates to a composition comprising a polypeptide as described above, a nucleic acid encoding a polypeptide as described above, a vector comprising such a nucleic acid and a composition comprising such a nucleic acid or a vector.

[0018] The invention also provides a recombinant host cell comprising a nucleic acid, a vector or a composition as described above.

[0019] Moreover, the invention relates to a method for producing a polypeptide as described above, comprising the steps of: culturing a recombinant host cell as described above, under conditions suitable for the production of the polypeptide, and recovering the polypeptide obtained, and optionally purifying the polypeptide.

[0020] In addition, the invention relates to a method of using a polypeptide as described above for converting glucose to fructose.

[0021] The invention also relates to a method for improving the glucose to fructose conversion in the presence of xylose by a polypeptide comprising an amino acid sequence that is at least 90% identical to the amino acid according to SEQ ID NO: 1 or SEQ ID NO: 2, the method comprising the step of altering the amino acid at a position corresponding to position 104 in SEQ ID NO: 1 or SEQ ID NO: 2 to a tiny amino acid residue.

[0022] Exemplified herein are improved glucose isomerases comprising an amino acid sequence according to SEQ ID NO: 1, wherein single amino acid substitutions have been made in order to arrive at glucose isomerases P104G, P104S, P104A and P104C. This annotation is used herein to indicate a replacement of the amino acid residue Proline (P), corresponding to position 104 of SEQ ID NO: 1, with either one of the residues G (glycine), S (serine), A (alanine) or C (cysteine), thereby obtaining the polypeptides according to SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5 and SEQ ID NO: 6, respectively.

[0023] Also exemplified are similar substitutions in a glucose isomerase comprising the amino acid sequence according to SEQ ID NO: 2, wherein glucose isomerases P104G, P104S,

P104A and P104C are obtained, the amino acid sequences of which are represented by SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9 and SEQ ID NO: 10 respectively.

[0024] DNA sequences encoding amino acid sequences SEQ ID NO: 1 through SEQ ID NO: 10 are represented by SEQ ID NO: 11 through SEQ ID NO: 20.

Legend to the figure

[0025] Figure 1: Diagram showing the relative glucose isomerase activity of polypeptides comprising the amino acid sequence according to SEQ ID NO: 1 (WT, annotated as GI1), SEQ ID NO: 2 (WT, annotated as GI2) and SEQ ID NO: 3 (GI1/P104G), SEQ ID NO: 4 (GI1/P104S), SEQ ID NO: 5 (GI1/P104A), SEQ ID NO: 6 (GI1/P104C), SEQ ID NO: 7 (GI2/P104G), SEQ ID NO: 8 (GI2/P104S), SEQ ID NO: 9 (GI2/P104A), SEQ ID NO: 10 (GI2/P104C) on various substrates, see example 4.

[0026] Glucose isomerase activity was determined using a substrate containing 200 mM glucose (labelled "Glucose") and compared to a substrate containing 200 mM Glucose and 20 mM xylose (labelled "Glucose+Xylose", see Example 4) as well as a hardwood hydrolysate comprising 42 mM xylose (see example 6).

Detailed description of the invention

[0027] In enzymology, a glucose isomerase (EC 5.3.1.5) is an enzyme that catalyzes the interconversion of D-glucose and D-fructose. This enzyme belongs to the family of isomerases, specifically those intramolecular oxidoreductases interconverting aldoses and ketoses. The glucose isomerase has now been observed in nearly a hundred species of bacteria. The systematic name of this enzyme class is D-xylose aldose-ketose-isomerase. Other names in common use include D-xylose isomerase, D-xylose ketoisomerase, and D-xylose ketol-isomerase. In industry, these enzymes are mostly referred to as glucose-isomerases due to their industrial use to produce fructose from glucose. However, xylose is a preferred substrate for these enzymes and the presence of xylose is highly inhibiting glucose isomerization.

[0028] The commercially available glucose isomerase enzymes have been used successfully in the production of high fructose corn syrup (HFCS) from starch, but they are not suited for the isomerisation of glucose obtained from lignocellulose material. Such lignocellulose derived glucose is characterized by the presence of lignin and other sugars derived from hemicelluloses including xylose.

[0029] Lignocellulosic biomass is the most abundantly available carbon source on earth for the production of biofuels, mainly bio-ethanol and potentially also for the production of bio-based materials such as polymers and plastics. It is composed of carbohydrate polymers (cellulose,

hemicellulose), and an aromatic polymer (lignin). Cellulose consists of linear glucose polymers, whereas hemicellulose is a branched heterogeneous polymer consisting of various 6- and 5-carbon sugars depending on the biological species.

[0030] Hemicellulose derived from softwood contains mostly glucomannans, while hemicellulose derived from hardwood contains mostly glucuronoxylans. Thus hydrolysate of hardwood usually contains a considerable amount of xylose (Sjostrom, E., Wood Chemistry. Fundamentals and Applications. Second edition ed. 1993, San Diego: Academic press. 292.).

[0031] In our co-pending European Patent Application EP 16175234.0 it was shown that wild type glucose isomerases GI1 and GI2 are efficiently functioning in softwood hydrolysate, where they are resistant to lignin and other impurities. However, hardwood presents an additional hurdle, namely xylose. Xylose is a much-preferred substrate for these type of enzymes, and reaction being reversible the xylose is not spent during the course of the reaction and poses continuous inhibition.

[0032] We herein indeed confirmed that the glucose isomerases derived from Dictyoglomus thermophilum and Dictyoglomus turgidum according to SEQ ID NO: 1 and SEQ ID NO: 2 were suitable for the conversion of glucose to fructose in a solution containing glucose, however, they were strongly inhibited by the presence of xylose such as present in hardwood hydrolysates (figure 1).

[0033] We found that GI1 and GI2 activities were reduced in the presence of xylose (i.e. in a solution of glucose plus xylose), reductions to 10-15% of their activity in pure glucose were observed. In line with these findings, the activity of both glucose isomerases was reduced to less than 12% when a hydrolysate of hardwood was used as the substrate (Table 1, figure 1).

[0034] Hardwood comes from angiosperm - or flowering plants - such as birch, eucalyptus, oak, maple, or walnut, which are not monocots. Other examples of hardwood include but are not limited to alder, balsa, beech, hickory, mahogany, and teak.

[0035] Surprisingly, we found that a single mutation in the wild type glucose isomerases GI1 and GI2 at position 104 substituting the wild type amino acid Proline with a tiny amino acid such as glycine, alanine, serine, or cysteine, makes glucose isomerization much less inhibited by xylose. Moreover, it tremendously improved the activity of the glucose isomerase to a value between 150 and 250% (Table 1).

Table 1 Relative glucose isomerase activity of GI1 and GI2 in different substrates

	Glucose [% activity]	**Glucose + Xylose [% activity]**	**Hardwood [% activity]**
GI1	100	15	12
GI1/P104G	250	240	150
GI1/P104S	245	230	140
GI1/P104A	230	215	120

	Glucose [% activity]	Glucose + Xylose [% activity]	Hardwood [% activity]
GI1/P104C	150	120	90
GI2	95	12	10
GI2/P104G	240	230	130
GI2/P104S	230	220	110
GI2/P104A	220	200	100
GI2/P104C	145	100	80

[0036] In conclusion, wild type glucose isomerase activity of enzymes GI1 and GI2 in 200 mM glucose was greatly reduced in the presence of 20mM xylose. In contrast, glucose isomerase enzymes carrying one of the above mentioned mutations at position 104 were more active than the wild type enzymes in the presence of xylose (table 1, figure 1, example 4). Moreover, enzymes with mutations at position 104 showed 1.5 to 2.5 fold higher glucose isomerization activity as compared to wild type enzymes at the same enzyme dosage (figure 1).

[0037] Next, we tested the mutated enzymes in the isomerization of glucose in crude lignocellulosic hydrolysates of hardwood (example 6). Hardwood hydrolysate, containing higher levels of xylose than softwood, greatly inhibited the glucose isomerase activity of the wild type enzymes GI1 and GI2, but not that of the mutated enzymes. Mutated enzymes demonstrated a clear glucose isomerase activity in hardwood hydrolysates, making them exceptionally suited for the conversion of glucose into fructose in lignocellulose-derived material, in particular when the lignocellulose is derived from wood such as wood with a high xylose content, such as hardwood.

[0038] The term "mutated enzyme" as used herein refers to a glucose isomerase enzyme comprising an amino acid sequence according to SEQ ID NO: 1 or SEQ ID NO: 2, wherein the amino acid at a position corresponding to the position 104 of SEQ ID NO: 1, or SEQ ID NO: 2 (proline in SEQ ID NO: 1 or SEQ ID NO: 2) is replaced by a tiny amino acid.

[0039] The term "tiny amino acid" as used herein indicates an amino acid with a tiny side group. In other words, these amino acids are smaller than 110 Angstrom or 11 nanometer. http://www.imgt.org/IMGTeducation/Aide-memoire/_UK/aminoacids/IMGTclasses.html.
Preferred examples of tiny amino acids are amino acids G (glycine), A (alanine), S (serine) and C (cysteine).

[0040] Glucose isomerases (GIs) according to SEQ ID NO: 1 and SEQ ID NO: 2 are homologous sequences with an amino acid sequence identity of 98%. It may therefore be expected that closely related GIs, such as GIs with an amino acid sequence that is at least 90% such as 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or even 99% identical with either SEQ ID NO: 1 or SEQ ID NO: 2, will perform in the same way as GI1 and GI2 exemplified herein. Such close homologues may be obtained from natural sources or by directed

mutagenesis. The skilled person is well aware of materials and methods for obtaining such close homologues.

[0041] As used herein, the degree of identity between two or more amino acid sequences is equivalent to a function of the number of identical positions shared by the sequences (i.e., % identity = number of identical positions divided by the total number of aligned positions x 100), excluding gaps, which need to be introduced for optimal alignment of the two sequences, and overhangs. The comparison of sequences and determination of percent identity between two or more sequences can be accomplished using standard methods known in the art. For example, a freeware conventionally used for this purpose is "Align" tool at NCBI recourse http://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE_TYPE=BlastSearch&BLAST_SPEC=blast 2seq&LINK_LOC=align2seq

[0042] Introduction of a specific mutation in a recombinant gene is also among the routine skills of a molecular biologist. Specific guidance may be obtained from Methods in Molecular Biology Vol 182, "In vitro mutagenesis protocols", Eds Jeff Braman, Humana Press 2002. There are commercially available kits for performing site-directed mutagenesis (for example, QuikChange II XL Site-Directed Mutagenesis kit Agilent Technologies cat No 200521).

[0043] Hence, the invention relates to a polypeptide with glucose isomerase activity comprising an amino acid sequence that is at least 90% identical to the amino acid according to SEQ ID NO: 1 or SEQ ID NO: 2, wherein the polypeptide comprises a tiny amino acid residue at an amino acid position corresponding to position 104 in SEQ ID NO: 1 or SEQ ID NO:2.

[0044] The phrase "an amino acid position corresponding to position 104 in SEQ ID NO: 1 or SEQ ID NO: 2" is in itself sufficiently clear for the skilled person. In order to avoid any misunderstanding, the following is provided. The phrase "an amino acid position corresponding to position 104 in SEQ ID NO: 1 or SEQ ID NO: 2" is used herein to indicate a certain position in the amino acid sequence of the polypeptide with glucose isomerase activity. That certain position is to be determined by aligning the sequence of the polypeptide with glucose isomerase activity with the sequence of either SEQ ID NO: 1 or SEQ ID NO: 2 as described above. The amino acid position in the polypeptide with glucose isomerase that aligns with the amino acid at position 104 in SEQ ID NO: 1 or SEQ ID NO: 2 is then referred to as the amino acid position corresponding to position 104 in SEQ ID NO: 1 or SEQ ID NO: 2.

[0045] The invention also relates to a method for the interconversion of D-glucose and D-fructose in the presence of a glucose isomerase, wherein the D-glucose is derived from lignocellulose-containing biomass, and wherein the glucose isomerase is a glucose isomerase according to the invention.

[0046] The phrase "glucose derived from lignocellulose-containing material" is equivalent to the term "lignocellulose-derived glucose". Both are used herein to indicate that the glucose is contained in a solution comprising other sugars, in particular xylose, derived from the lignocellulosic material, such as lignocellulosic biomass. As such, the term is used to

distinguish the glucose from purified glucose, which does not contain xylose.

[0047] Mutated variants of GI1 and GI2 as disclosed herein and their homologues with at least 90% sequence identity provide advantageous results in comparison to other GIs in conditions wherein the substrate solution comprises xylose. Not only are the mutant enzymes resistant against the presence of xylose in a composition comprising glucose and xylose, they are also more active, even up to 2.5 times more active in converting glucose into fructose.

[0048] In other terms, the invention relates to a process for converting glucose into fructose comprising the steps of:

1. a) providing a solution or suspension comprising glucose and xylose
2. b) enzymatically converting the glucose to fructose in the presence of a glucose isomerase,
3. c) optionally purifying the fructose from the solution,

wherein the glucose isomerase comprises an amino acid sequence that is at least 90% identical with the sequence according to SEQ ID NO: 1 or SEQ ID NO: 2 and wherein the glucose isomerase comprises a tiny amino acid residue at an amino acid position corresponding to position 104 in SEQ ID NO: 1 or SEQ ID NO: 2.

[0049] The solution or suspension comprising glucose and xylose may advantageously be obtained by hydrolyzing a hardwood biomass. Such hydrolysis is advantageously performed enzymatically, for instance by employing a cellulase.

[0050] Advantageously, the pretreatment step comprises a steam explosion step and/or an acid pretreatment step.

[0051] All these steps are well known in the art and the skilled person is well aware of the metes and bounds of the terms used herein.

[0052] The invention may have particular advantages when the enzyme is used in a solution or suspension that contains xylose in a concentration that inhibits the activity of the wild type enzyme according to SEQ ID NO: 1 or SEQ ID NO: 2 for 10% or more, such as 20%, 30%, 40%, 50% or more, such as 60%, 70%, 80%, 90% or even more like 95% or more.

[0053] The polypeptides as described herein may be used in compositions containing several additional components, such as stabilizers, fillers, cell debris, culture medium etcetera. Hence, the invention provides a composition comprising a polypeptide as described herein.

[0054] Polypeptides as described herein may be obtained by expressing a recombinant DNA in a heterologous expression system. The term "heterologous expression system" or equivalent means a system for expressing a DNA sequence from one host organism in a recipient organism from a different species or genus than the host organism. The most prevalent recipients, known as heterologous expression systems, are chosen usually because they are

easy to transfer DNA into or because they allow for a simpler assessment of the protein's function. Heterologous expression systems are also preferably used because they allow the upscaling of the production of a protein encoded by the DNA sequence in an industrial process. Preferred recipient organisms for use as heterologous expression systems include bacterial, fungal and yeast organisms, such as for example Escherichia coli, Bacillus, Corynebacterium, Pseudomonas, Pichia pastoris, Saccharomyces cerevisiae, Yarrowia lipolytica, filamentous fungi and many more systems well known in the art.

[0055] The presently disclosed polypeptides or proteins may be fused to additional sequences, by attaching or inserting, including , but not limited to, affinity tags, facilitating protein purification (S-tag, maltose binding domain, chitin binding domain), domains or sequences assisting folding (such as thioredoxin domain, SUMO protein), sequences affecting protein localization (periplasmic localization signals etc), proteins bearing additional function, such as green fluorescent protein (GFP), or sequences representing another enzymatic activity. Other suitable fusion partners for the presently disclosed polypeptides are known to those skilled in the art.

[0056] The present invention also relates to polynucleotides encoding any of the glucose isomerase variants disclosed herein. Means and methods for cloning and isolating such polynucleotides are well known in the art.

[0057] Furthermore, the present invention relates to a vector comprising a polynucleotide according to the invention, optionally operably linked to one or more control sequences. Suitable control sequences are readily available in the art and include, but are not limited to, promoter, leader, polyadenylation, and signal sequences.

[0058] Glucose isomerase variants according to various embodiments of the present invention may be obtained by standard recombinant methods known in the art. Briefly, such a method may comprise the steps of: culturing a recombinant host cell as described above under conditions suitable for the production of the polypeptide, and recovering the polypeptide obtained. The polypeptide may then optionally be further purified.

[0059] A large number of vector-host systems known in the art may be used for recombinant production of the glucose isomerases as described herein. Possible vectors include, but are not limited to, plasmids or modified viruses which are maintained in the host cell as autonomous DNA molecule or integrated in genomic DNA. The vector system must be compatible with the host cell used as is well known in the art. Non-limiting examples of suitable host cells include bacteria (e.g. E.coli, bacilli), yeast (e.g. Pichia Pastoris, Saccharomyces Cerevisiae), fungi (e.g. filamentous fungi) insect cells (e.g. Sf9).

[0060] In yet other terms, the invention relates to a method for improving glucose isomerization, especially in the presence of xylose, of a polypeptide with glucose isomerase activity comprising an amino acid sequence that is at least 90% identical to the amino acid according to SEQ ID NO: 1 or SEQ ID NO: 2, the method comprising the step of altering the

amino acid at a position corresponding to position 104 in SEQ ID NO: 1 or SEQ ID NO: 2 to a tiny amino acid residue.

Examples

Example 1: Preparation of polypeptides according to SEQ ID NO: 1 - 10 and nucleotides according to SEQ ID NO: 11 - 20.

[0061] The DNA constructs according to SEQ ID NO: 11 and SEQ ID NO: 12 encoding the polypeptides according to SEQ ID NO: 1, SEQ ID NO: 2 were designed using codon frequencies optimized for expression in E. coli and commercially synthesized and cloned into a plasmid vector based on a standard pET28a+ plasmid. The plasmid vector contained an N-terminal nucleotide sequence encoding peptidyl-prolyl isomerase from Enterobacteriaceae (Protein databank accession number WP_000255997.1). The recombinant gene was expressed in Escherichia coli BL21(DE3) under the control of the T7-RNA-polymerase promoter (see Example 2). This resulted in expression of the recombinant proteins with an N-terminal tag. Nucleotide sequences according to SEQ ID NO: 13 - 20 encoding the glucose isomerases according to SEQ ID NO: 3 - 10 were ordered commercially and expressed in the same way as described herein.

Example 2: Heterologous expression of polypeptides with glucose isomerase activity.

[0062] Protein production was carried out in E.coli BL21(DE3) strain according to the plasmid manufacturer protocol available at http://richsingiser.com/4402/Novagen%20pET%20system%20manual.pdf . The incubation temperature for protein production was 30 degrees Celsius, which was found optimal for maximum yield of the active protein. Cells were lysed using lysis buffer (50 mM Tris-HCl pH7.4, 1% Triton X100, 1 mM CoCl2) and heated at 70 degrees Celsius for 30 min. The glucose isomerase activity was detected in the insoluble fraction only, and could be fully recovered by centrifugation. Thus, thermostable recombinant glucose isomerase was expressed in active insoluble form allowing reuse of the enzyme in several reaction batches. Mutations at position corresponding to position 104 of GI1 or GI-2 did not detectably effect the expression level of the recombinant proteins, they were essentially the same as the expression levels of the wild type enzymes comprising amino acid sequences according to SEQ ID NO: 1 or SEQ ID NO:2.

Example 3: Glucose isomerase activity assay.

[0063] Glucose isomerase activity (isomerization reaction rate) was determined by measuring fructose level in the reaction mixture according to the protocol described in Schenk and

Bisswanger, A microplate assay for D-xylose/D-glucose isomerase. Enzyme and Microbial Technology (Elsevier Science Inc, NY, 1998), V22, pp.721-723.

[0064] Measurement was performed in the linear stage of the reaction course (product accumulation is linear with time). Ten-microliter aliquots of the reaction mixture were taken and pipette into a 96-well plate, 40 ul of water was added resulting in 50 ul sample. In some cases, higher dilution of the reaction mixture with water was used to prepare 50 ul of the diluted sample to match the dynamic range of the method. 150 ml of a freshly prepared 1:1 mixture (v/v) of solution A (0.05% resorcinol in ethanol) and solution B (0.216 g $FeNH4(SO4)2 * 12 H2O$ in 1 l concentrated HCl) were added. For color development, the plate was incubated at 80°C for 40 min. The absorbance was measured with a microplate reader (Thermo) at 490 nm.

[0065] It has to be noted that the presence of xylose or its isomerization product xylulose does not affect the measurement of fructose by this method.

Example 4 Glucose isomerization activity and xylose inhibition of polypeptides comprising SEQ ID NO: 1 - 10 in pure glucose solution with and without xylose.

[0066] In this experiment, we compared wild type glucose isomerases: GI1 (SEQ ID NO: 1) and GI2 (SEQ ID NO: 2) to glucose isomerases mutants P104G, P104S, P104A and P104C of both these sequences.

[0067] Enzymatic activity of the wild type enzyme (GI1 and GI2) and mutants thereof was first determined in a glucose solution (200 mM Glucose, 10 mM MOPS pH 8.0, 1 mM $MgCl2$).

[0068] A parallel set of reactions had the same composition apart from additionally comprising 20 mM xylose. The enzyme dosage was selected so that during the reaction time of 1h (at 70 degrees C), the product formation remained linear in all reactions. All enzymes were used at the same dosage (in micrograms of recombinant protein per ml of reaction). Glucose isomerization activity was measured as described in Example 3.

[0069] The results are shown in table 1 and depicted in Figure 1. Glucose isomerase activity of the wild type enzyme without xylose in the reaction mixture was taken as 100%, and activities in the presence of xylose as well as mutated enzymes activities in the presence or absence of xylose were calculated as percentage of this value.

[0070] Mutated glucose isomerases according to the invention showed much higher glucose isomerization activity (150-250% of the wild type). In addition, it was found that glucose isomerases according to the invention showed little or no sensitivity to the presence of xylose.

Example 5 Preparation of lignocellulose hydrolysate

[0071] Wood chips, obtained from birch (hardwood), were submerged in 2% sulfuric acid at a dry matter content of 20% and subjected to a steam explosion pretreatment essentially as described in European patent application EP 2623607A1. The pretreated material in its entirety (without removing solubilized fractions of hemicellulose and lignin) was subjected to enzymatic hydrolysis using Cellic® CTec3 cellulase product from Novozymes. The hydrolysis was carried out under the manufacturer's recommended conditions (incubation for 72 h at 55 degrees Celsius, pH 5.5 at 10% solids content), remaining solids were removed from the hydrolysis mixture by centrifugation, the liquid fraction was then evaporated to approximately 100 g/L sugar concentration, and the pH was adjusted to 8 with sodium hydroxide. The resulting solution is referred to herein further as "hydrolysate" or "lignocellulose hydrolysate" and used for the isomerization reaction. The resulting hardwood hydrolysate contained the following sugar composition: 68 g/L glucose, 12 g/l xylose and less than 1g/L other sugars

Example 6 Glucose isomerization activity and xylose inhibition of polypeptides comprising SEQ ID NO: 1 - 10 in lignocellulose hydrolysates of hardwood.

[0072] In this experiment, we compared wild type glucose isomerases: GI1 (SEQ ID NO: 1) and GI2 (SEQ ID NO: 2) with mutant variants P104G, P104S, P104A and P104C of both glucose isomerases for isomerization of glucose to fructose in crude hydrolysate of hardwood.

[0073] For this experiment, the lignocellulose hydrolysates (see example 5) were diluted with water to an end concentration of 200 mM glucose (resulting in 42 mM xylose concentration) and brought to 10 mM MOPS pH 8.0 and 1 mM MgCl2. Enzymes were added to the reaction mixtures at the same dosages as in Example 4 and reactions carried out for 1 h at 70 degrees C. The results are shown in Figure 1. For each enzyme, activity in 200 mM glucose solution without xylose as described above was taken as 100% and activity in hardwood hydrolysate was plotted as a percentage of that.

[0074] Both wild type enzymes were strongly inhibited in the hardwood hydrolysate, like they were in the glucose plus xylose solution described above. However, mutant enzymes were showing higher activities in both the glucose plus xylose model solution as in the hardwood hydrolysate (table 1, figure 1).

SEQUENCE LISTING

[0075]

<110> MetGen OY

<120> NEW IMPROVED GLUCOSE ISOMERASES

<130> 375 WO

<160> 20

<170> PatentIn version 3.5

<210> 1
<211> 368
<212> PRT
<213> Dictyoglomus thermophilum

<400> 1

```
Met Pro Phe Val Asp His Arg Ala Gln Lys Ile Arg Arg Ser Lys Glu
1               5                  10                  15

Glu Leu Leu Lys His Met Gln Thr Phe Lys Leu Asp Leu Lys Phe Ser
            20                  25                  30

Val Gly Ile Trp Tyr Phe Thr Pro Gly Gly Gly Arg Phe His Glu Pro
        35                  40                  45

Tyr Val Glu Gln Lys Ser Ile Pro Glu Arg Ile Glu Met Ala Ala Glu
    50                  55                  60

Met Ala Lys Phe Gly Val Lys Gly Ile Glu Ala His Tyr Pro Ala Glu
65                  70                  75                  80

Val Asn Glu Glu Asn Leu His Leu Tyr Lys Gln Leu Glu Lys Glu Ala
                85                  90                  95

Gly Ile Arg Leu Val Ala Val Pro Leu Ser Leu Phe Tyr Asp Lys Ile
            100                 105                 110

Phe Glu Phe Gly Ser Leu Ser Asn Pro Tyr Glu Lys Tyr Arg Lys Val
        115                 120                 125

Ala Tyr Glu Arg Leu Val Asn Gly Leu Lys Leu Val Lys Glu Ala Asn
    130                 135                 140

Ala Asp Ile Cys Ile Ile Trp Pro Gly Ile Asp Gly Tyr Thr Tyr Ser
145                 150                 155                 160

Tyr Gly His Leu Tyr Tyr His Met Trp Asp Thr Phe Glu Glu Leu Val
                165                 170                 175

Ala Gln Ala Met Asp Glu Val Pro Gly Val Gln Val Ala Ile Glu Pro
            180                 185                 190

Lys Pro Tyr Glu Pro Ala Pro Asn Asn Ile Tyr Arg Thr Thr Ala Asp
        195                 200                 205

Gly Ile Leu Ala Ala Arg Asp Ile Glu Ala Arg Leu Lys Asn Pro Glu
    210                 215                 220

Asn Leu Lys Leu Leu Gln Glu Glu His Ala Leu Val Glu Leu Asn Pro
```

```
Asn Leu Lys Leu Leu Gln Glu Gly His Ala Leu Val Gly Leu Asn Pro
225                 230                 235                 240

Glu Val Gly His Val Arg Met Gly Phe Glu Asp Leu Pro Tyr Ala Tyr
                245                 250                 255

Ala Arg Val Ala Arg Glu Gly Arg Leu Phe His Thr His Trp Asn Ser
            260                 265                 270

Gln Pro Leu Gly Asn Tyr Asp Gln Asp Leu Asn Ile Gly Val Val Asp
        275                 280                 285

Trp Asp Ser Thr Glu Ala Leu Leu Tyr Thr Leu Lys Met Val Gly Tyr
    290                 295                 300

Gln Gly Tyr Phe Gly Ile Asp Ile Asn Pro Glu Arg Met Pro Val Ile
305                 310                 315                 320

Lys Ala Ile Glu Ile Asn Thr Lys Val Leu Gln Ile Met Asn Glu Arg
                325                 330                 335

Ile Glu Arg Leu Pro His Asp Arg Ile Ile Glu Cys Tyr Phe Asp Pro
            340                 345                 350

Glu Asn His Arg Gly Glu Leu Glu Leu Ile Leu Ala Glu Asn His Lys
        355                 360                 365
```

<210> 2

<211> 368

<212> PRT

<213> Dictyoglomus thermophilum

<400> 2

```
Met Pro Phe Val Asp His Arg Asn Gln Lys Ile Arg Arg Ser Lys Glu
1               5                   10                  15

Glu Leu Leu Lys His Met Gln Thr Phe Lys Leu Asp Leu Lys Phe Ser
            20                  25                  30

Val Gly Ile Trp Tyr Phe Thr Pro Gly Gly Gly Arg Phe His Glu Pro
        35                  40                  45

Tyr Val Glu Gln Lys Gly Ile Pro Glu Arg Ile Glu Met Ala Ala Glu
    50                  55                  60

Met Ala Lys Tyr Gly Val Lys Gly Ile Glu Ala His Tyr Pro Ala Glu
65                  70                  75                  80

Val Asn Glu Glu Asn Leu His Leu Tyr Lys Gln Leu Glu Lys Glu Thr
                85                  90                  95

Gly Ile Arg Leu Val Ala Val Pro Leu Ser Leu Phe Tyr Asp Lys Ile
            100                 105                 110

Phe Glu Phe Gly Ser Leu Ser Asn Pro Tyr Glu Lys Tyr Arg Lys Ile
        115                 120                 125
```

```
Ala Tyr Glu Arg Leu Val Asn Gly Leu Lys Leu Val Lys Glu Ala Asn
    130                 135                 140

Ala Asp Ile Cys Ile Ile Trp Pro Gly Ile Asp Gly Tyr Thr Tyr Ser
145                 150                 155                 160

Tyr Gly His Leu Tyr Tyr His Met Trp Asp Thr Phe Glu Glu Leu Val
                165                 170                 175

Ala Gln Ala Met Asp Glu Val Pro Gly Val Gln Val Ala Ile Glu Pro
            180                 185                 190

Lys Pro Tyr Glu Pro Ala Pro Asn Asn Ile Tyr Arg Thr Thr Ala Asp
        195                 200                 205

Gly Ile Leu Ala Ala Arg Asp Ile Glu Ala Arg Leu Lys Asn Pro Glu
    210                 215                 220

Asn Leu Lys Leu Leu Gln Glu Gly His Ala Leu Val Gly Leu Asn Pro
225                 230                 235                 240

Glu Val Gly His Val Arg Met Gly Phe Glu Asp Leu Pro Tyr Ala Tyr
                245                 250                 255

Ala Arg Val Ala Arg Glu Gly Arg Leu Phe His Thr His Trp Asn Ser
            260                 265                 270

Gln Pro Leu Gly Asn Tyr Asp Gln Asp Leu Asn Ile Gly Val Val Asp
        275                 280                 285

Trp Asp Ser Thr Glu Ala Leu Leu Tyr Thr Leu Lys Met Val Gly Tyr
    290                 295                 300

Gln Gly Tyr Phe Gly Ile Asp Ile Asn Pro Glu Arg Ile Pro Val Val
305                 310                 315                 320

Lys Ala Ile Glu Ile Asn Thr Lys Val Leu Gln Ile Met Asn Glu Arg
                325                 330                 335

Ile Glu Arg Leu Pro His Asp Arg Ile Ile Glu Cys Tyr Phe Asp Pro
            340                 345                 350

Glu Asn His Arg Gly Glu Leu Glu Leu Ile Leu Ala Glu Asn His Arg
        355                 360                 365
```

<210> 3

<211> 368

<212> PRT

<213> Dictyoglomus thermophilum

<400> 3

```
Met Pro Phe Val Asp His Arg Ala Gln Lys Ile Arg Arg Ser Lys Glu
1               5                   10                  15

Glu Leu Leu Lys His Met Gln Thr Phe Lys Leu Asp Leu Lys Phe Ser
            20                  25                  30
```

20 25 30

Val Gly Ile Trp Tyr Phe Thr Pro Gly Gly Gly Arg Phe His Glu Pro
35 40 45

Tyr Val Glu Gln Lys Ser Ile Pro Glu Arg Ile Glu Met Ala Ala Glu
50 55 60

Met Ala Lys Phe Gly Val Lys Gly Ile Glu Ala His Tyr Pro Ala Glu
65 70 75 80

Val Asn Glu Glu Asn Leu His Leu Tyr Lys Gln Leu Glu Lys Glu Ala
85 90 95

Gly Ile Arg Leu Val Ala Val Gly Leu Ser Leu Phe Tyr Asp Lys Ile
100 105 110

Phe Glu Phe Gly Ser Leu Ser Asn Pro Tyr Glu Lys Tyr Arg Lys Val
115 120 125

Ala Tyr Glu Arg Leu Val Asn Gly Leu Lys Leu Val Lys Glu Ala Asn
130 135 140

Ala Asp Ile Cys Ile Ile Trp Pro Gly Ile Asp Gly Tyr Thr Tyr Ser
145 150 155 160

Tyr Gly His Leu Tyr Tyr His Met Trp Asp Thr Phe Glu Glu Leu Val
165 170 175

Ala Gln Ala Met Asp Glu Val Pro Gly Val Gln Val Ala Ile Glu Pro
180 185 190

Lys Pro Tyr Glu Pro Ala Pro Asn Asn Ile Tyr Arg Thr Thr Ala Asp
195 200 205

Gly Ile Leu Ala Ala Arg Asp Ile Glu Ala Arg Leu Lys Asn Pro Glu
210 215 220

Asn Leu Lys Leu Leu Gln Glu Gly His Ala Leu Val Gly Leu Asn Pro
225 230 235 240

Glu Val Gly His Val Arg Met Gly Phe Glu Asp Leu Pro Tyr Ala Tyr
245 250 255

Ala Arg Val Ala Arg Glu Gly Arg Leu Phe His Thr His Trp Asn Ser
260 265 270

Gln Pro Leu Gly Asn Tyr Asp Gln Asp Leu Asn Ile Gly Val Val Asp
275 280 285

Trp Asp Ser Thr Glu Ala Leu Leu Tyr Thr Leu Lys Met Val Gly Tyr
290 295 300

Gln Gly Tyr Phe Gly Ile Asp Ile Asn Pro Glu Arg Met Pro Val Ile
305 310 315 320

```
Lys Ala Ile Glu Ile Asn Thr Lys Val Leu Gln Ile Met Asn Glu Arg
                325                 330                 335

Ile Glu Arg Leu Pro His Asp Arg Ile Ile Glu Cys Tyr Phe Asp Pro
            340                 345                 350

Glu Asn His Arg Gly Glu Leu Glu Leu Ile Leu Ala Glu Asn His Lys
        355                 360                 365
```

<210> 4

<211> 368

<212> PRT

<213> Dictyoglomus thermophilum

<400> 4

```
Met Pro Phe Val Asp His Arg Ala Gln Lys Ile Arg Arg Ser Lys Glu
1               5                   10                  15

Glu Leu Leu Lys His Met Gln Thr Phe Lys Leu Asp Leu Lys Phe Ser
            20                  25                  30

Val Gly Ile Trp Tyr Phe Thr Pro Gly Gly Gly Arg Phe His Glu Pro
        35                  40                  45

Tyr Val Glu Gln Lys Ser Ile Pro Glu Arg Ile Glu Met Ala Ala Glu
    50                  55                  60

Met Ala Lys Phe Gly Val Lys Gly Ile Glu Ala His Tyr Pro Ala Glu
65                  70                  75                  80

Val Asn Glu Glu Asn Leu His Leu Tyr Lys Gln Leu Glu Lys Glu Ala
                85                  90                  95

Gly Ile Arg Leu Val Ala Val Ser Leu Ser Leu Phe Tyr Asp Lys Ile
            100                 105                 110

Phe Glu Phe Gly Ser Leu Ser Asn Pro Tyr Glu Lys Tyr Arg Lys Val
        115                 120                 125

Ala Tyr Glu Arg Leu Val Asn Gly Leu Lys Leu Val Lys Glu Ala Asn
    130                 135                 140

Ala Asp Ile Cys Ile Ile Trp Pro Gly Ile Asp Gly Tyr Thr Tyr Ser
145                 150                 155                 160

Tyr Gly His Leu Tyr Tyr His Met Trp Asp Thr Phe Glu Glu Leu Val
                165                 170                 175

Ala Gln Ala Met Asp Glu Val Pro Gly Val Gln Val Ala Ile Glu Pro
            180                 185                 190

Lys Pro Tyr Glu Pro Ala Pro Asn Asn Ile Tyr Arg Thr Thr Ala Asp
        195                 200                 205

Gly Ile Leu Ala Ala Arg Asp Ile Glu Ala Arg Leu Lys Asn Pro Glu
    210                 215                 220
```

```
Asn Leu Lys Leu Leu Gln Glu Gly His Ala Leu Val Gly Leu Asn Pro
225                 230                 235                 240

Glu Val Gly His Val Arg Met Gly Phe Glu Asp Leu Pro Tyr Ala Tyr
                245                 250                 255

Ala Arg Val Ala Arg Glu Gly Arg Leu Phe His Thr His Trp Asn Ser
            260                 265                 270

Gln Pro Leu Gly Asn Tyr Asp Gln Asp Leu Asn Ile Gly Val Val Asp
        275                 280                 285

Trp Asp Ser Thr Glu Ala Leu Leu Tyr Thr Leu Lys Met Val Gly Tyr
    290                 295                 300

Gln Gly Tyr Phe Gly Ile Asp Ile Asn Pro Glu Arg Met Pro Val Ile
305                 310                 315                 320

Lys Ala Ile Glu Ile Asn Thr Lys Val Leu Gln Ile Met Asn Glu Arg
                325                 330                 335

Ile Glu Arg Leu Pro His Asp Arg Ile Ile Glu Cys Tyr Phe Asp Pro
            340                 345                 350

Glu Asn His Arg Gly Glu Leu Glu Leu Ile Leu Ala Glu Asn His Lys
        355                 360                 365
```

<210> 5

<211> 368

<212> PRT

<213> Dictyoglomus thermophilum

<400> 5

```
Met Pro Phe Val Asp His Arg Ala Gln Lys Ile Arg Arg Ser Lys Glu
1               5                   10                  15

Glu Leu Leu Lys His Met Gln Thr Phe Lys Leu Asp Leu Lys Phe Ser
            20                  25                  30

Val Gly Ile Trp Tyr Phe Thr Pro Gly Gly Gly Arg Phe His Glu Pro
        35                  40                  45

Tyr Val Glu Gln Lys Ser Ile Pro Glu Arg Ile Glu Met Ala Ala Glu
    50                  55                  60

Met Ala Lys Phe Gly Val Lys Gly Ile Glu Ala His Tyr Pro Ala Glu
65                  70                  75                  80

Val Asn Glu Glu Asn Leu His Leu Tyr Lys Gln Leu Glu Lys Glu Ala
                85                  90                  95

Gly Ile Arg Leu Val Ala Val Ala Leu Ser Leu Phe Tyr Asp Lys Ile
            100                 105                 110

Phe Glu Phe Gly Ser Leu Ser Asn Pro Tyr Glu Lys Tyr Arg Lys Val
```

```
        115                 120                 125

Ala Tyr Glu Arg Leu Val Asn Gly Leu Lys Leu Val Lys Glu Ala Asn
    130                 135                 140

Ala Asp Ile Cys Ile Ile Trp Pro Gly Ile Asp Gly Tyr Thr Tyr Ser
145                 150                 155                 160

Tyr Gly His Leu Tyr Tyr His Met Trp Asp Thr Phe Glu Glu Leu Val
                165                 170                 175

Ala Gln Ala Met Asp Glu Val Pro Gly Val Gln Val Ala Ile Glu Pro
            180                 185                 190

Lys Pro Tyr Glu Pro Ala Pro Asn Asn Ile Tyr Arg Thr Thr Ala Asp
        195                 200                 205

Gly Ile Leu Ala Ala Arg Asp Ile Glu Ala Arg Leu Lys Asn Pro Glu
    210                 215                 220

Asn Leu Lys Leu Leu Gln Glu Gly His Ala Leu Val Gly Leu Asn Pro
225                 230                 235                 240

Glu Val Gly His Val Arg Met Gly Phe Glu Asp Leu Pro Tyr Ala Tyr
                245                 250                 255

Ala Arg Val Ala Arg Glu Gly Arg Leu Phe His Thr His Trp Asn Ser
            260                 265                 270

Gln Pro Leu Gly Asn Tyr Asp Gln Asp Leu Asn Ile Gly Val Val Asp
        275                 280                 285

Trp Asp Ser Thr Glu Ala Leu Leu Tyr Thr Leu Lys Met Val Gly Tyr
    290                 295                 300

Gln Gly Tyr Phe Gly Ile Asp Ile Asn Pro Glu Arg Met Pro Val Ile
305                 310                 315                 320

Lys Ala Ile Glu Ile Asn Thr Lys Val Leu Gln Ile Met Asn Glu Arg
                325                 330                 335

Ile Glu Arg Leu Pro His Asp Arg Ile Ile Glu Cys Tyr Phe Asp Pro
            340                 345                 350

Glu Asn His Arg Gly Glu Leu Glu Leu Ile Leu Ala Glu Asn His Lys

355                 360                 365
```

<210> 6

<211> 368

<212> PRT

<213> Dictyoglomus thermophilum

<400> 6

```
Met Pro Phe Val Asp His Arg Ala Gln Lys Ile Arg Arg Ser Lys Glu
```

```
1               5                  10                  15

Glu Leu Leu Lys His Met Gln Thr Phe Lys Leu Asp Leu Lys Phe Ser
            20                  25                  30

Val Gly Ile Trp Tyr Phe Thr Pro Gly Gly Gly Arg Phe His Glu Pro
        35                  40                  45

Tyr Val Glu Gln Lys Ser Ile Pro Glu Arg Ile Glu Met Ala Ala Glu
    50                  55                  60

Met Ala Lys Phe Gly Val Lys Gly Ile Glu Ala His Tyr Pro Ala Glu
65                  70                  75                  80

Val Asn Glu Glu Asn Leu His Leu Tyr Lys Gln Leu Glu Lys Glu Ala
                85                  90                  95

Gly Ile Arg Leu Val Ala Val Cys Leu Ser Leu Phe Tyr Asp Lys Ile
            100                 105                 110

Phe Glu Phe Gly Ser Leu Ser Asn Pro Tyr Glu Lys Tyr Arg Lys Val
        115                 120                 125

Ala Tyr Glu Arg Leu Val Asn Gly Leu Lys Leu Val Lys Glu Ala Asn
    130                 135                 140

Ala Asp Ile Cys Ile Ile Trp Pro Gly Ile Asp Gly Tyr Thr Tyr Ser
145                 150                 155                 160

Tyr Gly His Leu Tyr Tyr His Met Trp Asp Thr Phe Glu Glu Leu Val
                165                 170                 175

Ala Gln Ala Met Asp Glu Val Pro Gly Val Gln Val Ala Ile Glu Pro
            180                 185                 190

Lys Pro Tyr Glu Pro Ala Pro Asn Asn Ile Tyr Arg Thr Thr Ala Asp
        195                 200                 205

Gly Ile Leu Ala Ala Arg Asp Ile Glu Ala Arg Leu Lys Asn Pro Glu

    210                 215                 220

Asn Leu Lys Leu Leu Gln Glu Gly His Ala Leu Val Gly Leu Asn Pro
225                 230                 235                 240

Glu Val Gly His Val Arg Met Gly Phe Glu Asp Leu Pro Tyr Ala Tyr
                245                 250                 255

Ala Arg Val Ala Arg Glu Gly Arg Leu Phe His Thr His Trp Asn Ser
            260                 265                 270

Gln Pro Leu Gly Asn Tyr Asp Gln Asp Leu Asn Ile Gly Val Val Asp
        275                 280                 285

Trp Asp Ser Thr Glu Ala Leu Leu Tyr Thr Leu Lys Met Val Gly Tyr
```

```
    290                 295                 300

Gln Gly Tyr Phe Gly Ile Asp Ile Asn Pro Glu Arg Met Pro Val Ile
305                 310                 315                 320

Lys Ala Ile Glu Ile Asn Thr Lys Val Leu Gln Ile Met Asn Glu Arg
                325                 330                 335

Ile Glu Arg Leu Pro His Asp Arg Ile Ile Glu Cys Tyr Phe Asp Pro
            340                 345                 350

Glu Asn His Arg Gly Glu Leu Glu Leu Ile Leu Ala Glu Asn His Lys
        355                 360                 365
```

<210> 7

<211> 368

<212> PRT

<213> Dictyoglomus thermophilum

<400> 7

```
Met Pro Phe Val Asp His Arg Asn Gln Lys Ile Arg Arg Ser Lys Glu
1               5                   10                  15

Glu Leu Leu Lys His Met Gln Thr Phe Lys Leu Asp Leu Lys Phe Ser
            20                  25                  30

Val Gly Ile Trp Tyr Phe Thr Pro Gly Gly Gly Arg Phe His Glu Pro
        35                  40                  45

Tyr Val Glu Gln Lys Gly Ile Pro Glu Arg Ile Glu Met Ala Ala Glu
    50                  55                  60

Met Ala Lys Tyr Gly Val Lys Gly Ile Glu Ala His Tyr Pro Ala Glu

65                  70                  75                  80

Val Asn Glu Glu Asn Leu His Leu Tyr Lys Gln Leu Glu Lys Glu Thr
                85                  90                  95

Gly Ile Arg Leu Val Ala Val Gly Leu Ser Leu Phe Tyr Asp Lys Ile
            100                 105                 110

Phe Glu Phe Gly Ser Leu Ser Asn Pro Tyr Glu Lys Tyr Arg Lys Ile
        115                 120                 125

Ala Tyr Glu Arg Leu Val Asn Gly Leu Lys Leu Val Lys Glu Ala Asn
    130                 135                 140

Ala Asp Ile Cys Ile Ile Trp Pro Gly Ile Asp Gly Tyr Thr Tyr Ser
145                 150                 155                 160

Tyr Gly His Leu Tyr Tyr His Met Trp Asp Thr Phe Glu Glu Leu Val
                165                 170                 175

Ala Gln Ala Met Asp Glu Val Pro Gly Val Gln Val Ala Ile Glu Pro
```

```
            180                 185                 190

Lys Pro Tyr Glu Pro Ala Pro Asn Asn Ile Tyr Arg Thr Thr Ala Asp
        195                 200                 205

Gly Ile Leu Ala Ala Arg Asp Ile Glu Ala Arg Leu Lys Asn Pro Glu
    210                 215                 220

Asn Leu Lys Leu Leu Gln Glu Gly His Ala Leu Val Gly Leu Asn Pro
225                 230                 235                 240

Glu Val Gly His Val Arg Met Gly Phe Glu Asp Leu Pro Tyr Ala Tyr
                245                 250                 255

Ala Arg Val Ala Arg Glu Gly Arg Leu Phe His Thr His Trp Asn Ser
            260                 265                 270

Gln Pro Leu Gly Asn Tyr Asp Gln Asp Leu Asn Ile Gly Val Val Asp
        275                 280                 285

Trp Asp Ser Thr Glu Ala Leu Leu Tyr Thr Leu Lys Met Val Gly Tyr
    290                 295                 300

Gln Gly Tyr Phe Gly Ile Asp Ile Asn Pro Glu Arg Ile Pro Val Val
305                 310                 315                 320

Lys Ala Ile Glu Ile Asn Thr Lys Val Leu Gln Ile Met Asn Glu Arg
                325                 330                 335

Ile Glu Arg Leu Pro His Asp Arg Ile Ile Glu Cys Tyr Phe Asp Pro
            340                 345                 350

Glu Asn His Arg Gly Glu Leu Glu Leu Ile Leu Ala Glu Asn His Arg
        355                 360                 365
```

<210> 8

<211> 368

<212> PRT

<213> Dictyoglomus thermophilum

<400> 8

```
Met Pro Phe Val Asp His Arg Asn Gln Lys Ile Arg Arg Ser Lys Glu
1               5                   10                  15

Glu Leu Leu Lys His Met Gln Thr Phe Lys Leu Asp Leu Lys Phe Ser
            20                  25                  30

Val Gly Ile Trp Tyr Phe Thr Pro Gly Gly Gly Arg Phe His Glu Pro
        35                  40                  45

Tyr Val Glu Gln Lys Gly Ile Pro Glu Arg Ile Glu Met Ala Ala Glu
    50                  55                  60

Met Ala Lys Tyr Gly Val Lys Gly Ile Glu Ala His Tyr Pro Ala Glu
65                  70                  75                  80
```

```
Val Asn Glu Glu Asn Leu His Leu Tyr Lys Gln Leu Glu Lys Glu Thr
                85                  90                  95

Gly Ile Arg Leu Val Ala Val Ser Leu Ser Leu Phe Tyr Asp Lys Ile
            100                 105                 110

Phe Glu Phe Gly Ser Leu Ser Asn Pro Tyr Glu Lys Tyr Arg Lys Ile
        115                 120                 125

Ala Tyr Glu Arg Leu Val Asn Gly Leu Lys Leu Val Lys Glu Ala Asn
    130                 135                 140

Ala Asp Ile Cys Ile Ile Trp Pro Gly Ile Asp Gly Tyr Thr Tyr Ser
145                 150                 155                 160

Tyr Gly His Leu Tyr Tyr His Met Trp Asp Thr Phe Glu Glu Leu Val
                165                 170                 175

Ala Gln Ala Met Asp Glu Val Pro Gly Val Gln Val Ala Ile Glu Pro
            180                 185                 190

Lys Pro Tyr Glu Pro Ala Pro Asn Asn Ile Tyr Arg Thr Thr Ala Asp
        195                 200                 205

Gly Ile Leu Ala Ala Arg Asp Ile Glu Ala Arg Leu Lys Asn Pro Glu
    210                 215                 220

Asn Leu Lys Leu Leu Gln Glu Gly His Ala Leu Val Gly Leu Asn Pro
225                 230                 235                 240

Glu Val Gly His Val Arg Met Gly Phe Glu Asp Leu Pro Tyr Ala Tyr
                245                 250                 255

Ala Arg Val Ala Arg Glu Gly Arg Leu Phe His Thr His Trp Asn Ser
            260                 265                 270

Gln Pro Leu Gly Asn Tyr Asp Gln Asp Leu Asn Ile Gly Val Val Asp
        275                 280                 285

Trp Asp Ser Thr Glu Ala Leu Leu Tyr Thr Leu Lys Met Val Gly Tyr
    290                 295                 300

Gln Gly Tyr Phe Gly Ile Asp Ile Asn Pro Glu Arg Ile Pro Val Val
305                 310                 315                 320

Lys Ala Ile Glu Ile Asn Thr Lys Val Leu Gln Ile Met Asn Glu Arg
                325                 330                 335

Ile Glu Arg Leu Pro His Asp Arg Ile Ile Glu Cys Tyr Phe Asp Pro
            340                 345                 350

Glu Asn His Arg Gly Glu Leu Glu Leu Ile Leu Ala Glu Asn His Arg
        355                 360                 365
```

<210> 9

<211> 368

<212> PRT

<213> Dictyoglomus thermophilum

<400> 9

```
Met Pro Phe Val Asp His Arg Asn Gln Lys Ile Arg Arg Ser Lys Glu
1               5                  10                  15

Glu Leu Leu Lys His Met Gln Thr Phe Lys Leu Asp Leu Lys Phe Ser
            20                  25                  30

Val Gly Ile Trp Tyr Phe Thr Pro Gly Gly Gly Arg Phe His Glu Pro
        35                  40                  45

Tyr Val Glu Gln Lys Gly Ile Pro Glu Arg Ile Glu Met Ala Ala Glu
    50                  55                  60

Met Ala Lys Tyr Gly Val Lys Gly Ile Glu Ala His Tyr Pro Ala Glu
65                  70                  75                  80

Val Asn Glu Glu Asn Leu His Leu Tyr Lys Gln Leu Glu Lys Glu Thr
                85                  90                  95

Gly Ile Arg Leu Val Ala Val Ala Leu Ser Leu Phe Tyr Asp Lys Ile
            100                 105                 110

Phe Glu Phe Gly Ser Leu Ser Asn Pro Tyr Glu Lys Tyr Arg Lys Ile
        115                 120                 125

Ala Tyr Glu Arg Leu Val Asn Gly Leu Lys Leu Val Lys Glu Ala Asn
    130                 135                 140

Ala Asp Ile Cys Ile Ile Trp Pro Gly Ile Asp Gly Tyr Thr Tyr Ser
145                 150                 155                 160

Tyr Gly His Leu Tyr Tyr His Met Trp Asp Thr Phe Glu Glu Leu Val
                165                 170                 175

Ala Gln Ala Met Asp Glu Val Pro Gly Val Gln Val Ala Ile Glu Pro
            180                 185                 190

Lys Pro Tyr Glu Pro Ala Pro Asn Asn Ile Tyr Arg Thr Thr Ala Asp
        195                 200                 205

Gly Ile Leu Ala Ala Arg Asp Ile Glu Ala Arg Leu Lys Asn Pro Glu
    210                 215                 220

Asn Leu Lys Leu Leu Gln Glu Gly His Ala Leu Val Gly Leu Asn Pro
225                 230                 235                 240

Glu Val Gly His Val Arg Met Gly Phe Glu Asp Leu Pro Tyr Ala Tyr
                245                 250                 255

Ala Arg Val Ala Arg Glu Gly Arg Leu Phe His Thr His Trp Asn Ser
            260                 265                 270

-- - - -- - - - -- - - - -- -- -- - -- - -
```

```
Gln Pro Leu Gly Asn Tyr Asp Gln Asp Leu Asn Ile Gly Val Val Asp
        275                 280                 285

Trp Asp Ser Thr Glu Ala Leu Leu Tyr Thr Leu Lys Met Val Gly Tyr
    290                 295                 300

Gln Gly Tyr Phe Gly Ile Asp Ile Asn Pro Glu Arg Ile Pro Val Val
305                 310                 315                 320

Lys Ala Ile Glu Ile Asn Thr Lys Val Leu Gln Ile Met Asn Glu Arg
                325                 330                 335

Ile Glu Arg Leu Pro His Asp Arg Ile Ile Glu Cys Tyr Phe Asp Pro
            340                 345                 350

Glu Asn His Arg Gly Glu Leu Glu Leu Ile Leu Ala Glu Asn His Arg
        355                 360                 365
```

<210> 10

<211> 368

<212> PRT

<213> Dictyoglomus thermophilum

<400> 10

```
Met Pro Phe Val Asp His Arg Asn Gln Lys Ile Arg Arg Ser Lys Glu
1               5                   10                  15

Glu Leu Leu Lys His Met Gln Thr Phe Lys Leu Asp Leu Lys Phe Ser
            20                  25                  30

Val Gly Ile Trp Tyr Phe Thr Pro Gly Gly Gly Arg Phe His Glu Pro
        35                  40                  45

Tyr Val Glu Gln Lys Gly Ile Pro Glu Arg Ile Glu Met Ala Ala Glu
    50                  55                  60

Met Ala Lys Tyr Gly Val Lys Gly Ile Glu Ala His Tyr Pro Ala Glu
65                  70                  75                  80

Val Asn Glu Glu Asn Leu His Leu Tyr Lys Gln Leu Glu Lys Glu Thr
                85                  90                  95

Gly Ile Arg Leu Val Ala Val Cys Leu Ser Leu Phe Tyr Asp Lys Ile
            100                 105                 110

Phe Glu Phe Gly Ser Leu Ser Asn Pro Tyr Glu Lys Tyr Arg Lys Ile
        115                 120                 125

Ala Tyr Glu Arg Leu Val Asn Gly Leu Lys Leu Val Lys Glu Ala Asn
    130                 135                 140

Ala Asp Ile Cys Ile Ile Trp Pro Gly Ile Asp Gly Tyr Thr Tyr Ser
145                 150                 155                 160

Tyr Gly His Leu Tyr Tyr His Met Trp Asp Thr Phe Glu Glu Leu Val
                165                 170                 175
```

```
Ala Gln Ala Met Asp Glu Val Pro Gly Val Gln Val Ala Ile Glu Pro
            180                 185                 190

Lys Pro Tyr Glu Pro Ala Pro Asn Asn Ile Tyr Arg Thr Thr Ala Asp
        195                 200                 205

Gly Ile Leu Ala Ala Arg Asp Ile Glu Ala Arg Leu Lys Asn Pro Glu
    210                 215                 220

Asn Leu Lys Leu Leu Gln Glu Gly His Ala Leu Val Gly Leu Asn Pro
225                 230                 235                 240

Glu Val Gly His Val Arg Met Gly Phe Glu Asp Leu Pro Tyr Ala Tyr
                245                 250                 255

Ala Arg Val Ala Arg Glu Gly Arg Leu Phe His Thr His Trp Asn Ser
            260                 265                 270

Gln Pro Leu Gly Asn Tyr Asp Gln Asp Leu Asn Ile Gly Val Val Asp
        275                 280                 285

Trp Asp Ser Thr Glu Ala Leu Leu Tyr Thr Leu Lys Met Val Gly Tyr
    290                 295                 300

Gln Gly Tyr Phe Gly Ile Asp Ile Asn Pro Glu Arg Ile Pro Val Val
305                 310                 315                 320

Lys Ala Ile Glu Ile Asn Thr Lys Val Leu Gln Ile Met Asn Glu Arg
                325                 330                 335

Ile Glu Arg Leu Pro His Asp Arg Ile Ile Glu Cys Tyr Phe Asp Pro
            340                 345                 350

Glu Asn His Arg Gly Glu Leu Glu Leu Ile Leu Ala Glu Asn His Arg
        355                 360                 365
```

<210> 11

<211> 1104

<212> DNA

<213> Dictyoglomus thermophilum

<400> 11

```
atgccgtttg ttgatcatcg tgcacagaaa attcgtcgca gcaaagaaga actgctgaaa      60

catatgcaga ccttcaaact ggatctgaaa tttagcgtgg gcatctggta ttttacaccg     120

ggtggtggtc gttttcatga accgtatgtt gaacagaaaa gcattccgga acgtattgaa     180

atggcagcag aaatggcaaa atttggcgtg aaaggtattg aagcacatta tccggctgaa     240

gtgaatgaag aaaatctgca cctgtataaa cagctggaaa aagaagcagg tattcgtctg     300

gttgcagttc cgctgagcct gttttatgat aaaatctttg aatttggcag cctgagcaac     360

ccgtatgaaa aatatcgtaa agttgcctat gaacgcctgg tgaatggtct gaaactggtt     420

aaagaagcaa acgccgatat ttgcattatt tggcctggta ttgatggcta tacctatagc     480
```

tatggtcacc tgtattatca catgtgggat acctttgaag aactggttgc acaggcaatg 540

gatgaagttc cgggtgttca ggttgcaatt gaaccgaaac cgtatgaacc ggcaccgaat 600

aacatttatc gtaccaccgc agatggtatt ctggcagcac gtgatattga agcgcgtctg 660

aaaaatccgg aaaacctgaa actgctgcaa gaaggtcacg cactggttgg tctgaatccg 720

gaagttggtc atgttcgtat gggttttgaa gatctgccgt atgcatatgc ccgtgttgca 780

cgtgaaggtc gtctgtttca tacccattgg aatagccagc cgctgggtaa ttatgatcag 840

gatctgaata ttggtgtggt ggattgggat agcaccgaag cactgctgta taccctgaaa 900

atggttggtt atcagggcta ttttggcatc gatatcaatc cggaacgcat gccggttatt 960

aaagccattg aaattaacac caaagtgctg cagattatga acgaacgcat tgaacgtctg 1020

ccgcatgatc gtattattga gtgttatttt gaccctgaga atcatcgtgg tgaactggaa 1080

ctgattctgg ccgaaaatca taaa 1104

<210> 12
<211> 1104
<212> DNA
<213> Dictyoglomus thermophilum

<400> 12
atgccgtttg tggatcatcg taatcagaaa attcgtcgca gcaaagaaga actgctgaaa 60

cacatgcaga cctttaaact ggatctgaaa tttagcgttg gcatctggta ttttaccccct 120

ggtggtggtc gttttcatga accgtatgtt gaacagaaag gtattccgga acgtattgaa 180

atggcagcag aaatggcaaa atatggcgtt aaaggtatcg aagcacatta tccggctgaa 240

gtgaatgaag aaaatctgca cctgtataaa cagctggaaa aagaaaccgg tattcgtctg 300

gttgcagttc cgctgagcct gttttatgat aaaatctttg aatttggcag cctgagcaac 360

ccgtatgaaa aatatcgtaa aattgcctat gaacgcctgg tgaatggtct gaaactggtt 420

aaagaagcaa acgccgatat ttgcattatt tggcctggta ttgatggcta tacctatagc 480

tatggtcacc tgtattatca catgtgggat acctttgaag aactggttgc acaggcaatg 540

gatgaagttc cgggtgttca ggttgcaatt gaaccgaaac cgtatgaacc ggcaccgaat 600

aacatttatc gtaccaccgc agatggtatt ctggcagcac gtgatattga agcacgtctg 660

aaaaatccgg aaaacctgaa actgctgcaa gaaggtcacg cactggttgg tctgaatccg 720

gaagttggtc atgttcgtat gggttttgaa gatctgccgt atgcctatgc acgtgttgca 780

cgtgaaggtc gtctgtttca tacccattgg aatagccagc cgctgggtaa ttatgatcag 840

gatctgaata ttggtgtggt ggattgggat agcaccgaag cactgctgta taccctgaaa 900

atggttggtt atcagggcta ttttggcatc gatattaatc cggaacgcat tccggttgtt 960

aaagccattg aaattaacac caaagtgctg cagattatga acgaacgcat tgaacgtctg 1020

ccgcatgatc gtattattga gtgttatttt gaccccgaaa atcatcgtgg tgaactggaa 1080

ctgattctgg cggaaaacca tcgt 1104

<210> 13

<211> 1104
<212> DNA
<213> Dictyoglomus thermophilum

<400> 13
atgccgtttg ttgatcatcg tgcacagaaa attcgtcgca gcaaagaaga actgctgaaa 60

catatgcaga ccttcaaact ggatctgaaa tttagcgtgg gcatctggta ttttacaccg 120

ggtggtggtc gttttcatga accgtatgtt gaacagaaaa gcattccgga acgtattgaa 180

atggcagcag aaatggcaaa atttggcgtg aaaggtattg aagcacatta tccggctgaa 240

gtgaatgaag aaaatctgca cctgtataaa cagctggaaa aagaagcagg tattcgtctg 300

gttgcagttg gtctgagcct gttttatgat aaaatctttg aatttggcag cctgagcaac 360

ccgtatgaaa aatatcgtaa agttgcctat gaacgcctgg tgaatggtct gaaactggtt 420

aaagaagcaa acgccgatat ttgcattatt tggcctggta ttgatggcta tacctatagc 480

tatggtcacc tgtattatca catgtgggat acctttgaag aactggttgc acaggcaatg 540

gatgaagttc cgggtgttca ggttgcaatt gaaccgaaac cgtatgaacc ggcaccgaat 600

aacatttatc gtaccaccgc agatggtatt ctggcagcac gtgatattga agcgcgtctg 660

aaaaatccgg aaaacctgaa actgctgcaa gaaggtcacg cactggttgg tctgaatccg 720

gaagttggtc atgttcgtat gggttttgaa gatctgccgt atgcatatgc ccgtgttgca 780

cgtgaaggtc gtctgtttca tacccattgg aatagccagc cgctgggtaa ttatgatcag 840

gatctgaata ttggtgtggt ggattgggat agcaccgaag cactgctgta taccctgaaa 900

atggttggtt atcagggcta ttttggcatc gatatcaatc cggaacgcat gccggttatt 960

aaagccattg aaattaacac caaagtgctg cagattatga acgaacgcat tgaacgtctg 1020

ccgcatgatc gtattattga gtgttatttt gaccctgaga atcatcgtgg tgaactggaa 1080

ctgattctgg ccgaaaatca taaa 1104

<210> 14
<211> 1104
<212> DNA
<213> Dictyoglomus thermophilum

<400> 14
atgccgtttg ttgatcatcg tgcacagaaa attcgtcgca gcaaagaaga actgctgaaa 60

catatgcaga ccttcaaact ggatctgaaa tttagcgtgg gcatctggta ttttacaccg 120

ggtggtggtc gttttcatga accgtatgtt gaacagaaaa gcattccgga acgtattgaa 180

atggcagcag aaatggcaaa atttggcgtg aaaggtattg aagcacatta tccggctgaa 240

gtgaatgaag aaaatctgca cctgtataaa cagctggaaa aagaagcagg tattcgtctg 300

gttgcagttt ctctgagcct gttttatgat aaaatctttg aatttggcag cctgagcaac 360

ccgtatgaaa aatatcgtaa agttgcctat gaacgcctgg tgaatggtct gaaactggtt 420

aaagaagcaa acgccgatat ttgcattatt tggcctggta ttgatggcta tacctatagc 480

tatggtcacc tgtattatca catgtgggat acctttgaag aactggttgc acaggcaatg 540

gatgaagttc cgggtgttca ggttgcaatt gaaccgaaac cgtatgaacc ggcaccgaat 600

aacatttatc gtaccaccgc agatggtatt ctggcagcac gtgatattga agcgcgtctg 660

aaaaatccgg aaaacctgaa actgctgcaa gaaggtcacg cactggttgg tctgaatccg 720

gaagttggtc atgttcgtat gggttttgaa gatctgccgt atgcatatgc ccgtgttgca 780

cgtgaaggtc gtctgtttca tacccattgg aatagccagc cgctgggtaa ttatgatcag 840

gatctgaata ttggtgtggt ggattgggat agcaccgaag cactgctgta taccctgaaa 900

atggttggtt atcagggcta ttttggcatc gatatcaatc cggaacgcat gccggttatt 960

aaagccattg aaattaacac caaagtgctg cagattatga acgaacgcat tgaacgtctg 1020

ccgcatgatc gtattattga gtgttatttt gaccctgaga atcatcgtgg tgaactggaa 1080

ctgattctgg ccgaaaatca taaa 1104

<210> 15
<211> 1104
<212> DNA
<213> Dictyoglomus thermophilum

<400> 15
atgccgtttg ttgatcatcg tgcacagaaa attcgtcgca gcaaagaaga actgctgaaa 60

catatgcaga ccttcaaact ggatctgaaa tttagcgtgg gcatctggta ttttacaccg 120

ggtggtggtc gttttcatga accgtatgtt gaacagaaaa gcattccgga acgtattgaa 180

atggcagcag aaatggcaaa atttggcgtg aaaggtattg aagcacatta tccggctgaa 240

gtgaatgaag aaaatctgca cctgtataaa cagctggaaa aagaagcagg tattcgtctg 300

gttgcagttg ctctgagcct gttttatgat aaaatctttg aatttggcag cctgagcaac 360

ccgtatgaaa aatatcgtaa agttgcctat gaacgcctgg tgaatggtct gaaactggtt 420

aaagaagcaa acgccgatat ttgcattatt tggcctggta ttgatggcta tacctatagc 480

tatggtcacc tgtattatca catgtgggat acctttgaag aactggttgc acaggcaatg 540

gatgaagttc cgggtgttca ggttgcaatt gaaccgaaac cgtatgaacc ggcaccgaat 600

aacatttatc gtaccaccgc agatggtatt ctggcagcac gtgatattga agcgcgtctg 660

aaaaatccgg aaaacctgaa actgctgcaa gaaggtcacg cactggttgg tctgaatccg 720

gaagttggtc atgttcgtat gggttttgaa gatctgccgt atgcatatgc ccgtgttgca 780

cgtgaaggtc gtctgtttca tacccattgg aatagccagc cgctgggtaa ttatgatcag 840

gatctgaata ttggtgtggt ggattgggat agcaccgaag cactgctgta taccctgaaa 900

atggttggtt atcagggcta ttttggcatc gatatcaatc cggaacgcat gccggttatt 960

aaagccattg aaattaacac caaagtgctg cagattatga acgaacgcat tgaacgtctg 1020

ccgcatgatc gtattattga gtgttatttt gaccctgaga atcatcgtgg tgaactggaa 1080

ctgattctgg ccgaaaatca taaa 1104

<210> 16
<211> 1104
<212> DNA

<213> Dictyoglomus thermophilum

<400> 16

atgccgtttg ttgatcatcg tgcacagaaa attcgtcgca gcaaagaaga actgctgaaa 60

catatgcaga ccttcaaact ggatctgaaa tttagcgtgg gcatctggta ttttacaccg 120

ggtggtggtc gttttcatga accgtatgtt gaacagaaaa gcattccgga acgtattgaa 180

atggcagcag aaatggcaaa atttggcgtg aaaggtattg aagcacatta tccggctgaa 240

gtgaatgaag aaaatctgca cctgtataaa cagctggaaa aagaagcagg tattcgtctg 300

gttgcagttt gcctgagcct gttttatgat aaaatctttg aatttggcag cctgagcaac 360

ccgtatgaaa aatatcgtaa agttgcctat gaacgcctgg tgaatggtct gaaactggtt 420

aaagaagcaa acgccgatat ttgcattatt tggcctggta ttgatggcta tacctatagc 480

tatggtcacc tgtattatca catgtgggat acctttgaag aactggttgc acaggcaatg 540

gatgaagttc cgggtgttca ggttgcaatt gaaccgaaac cgtatgaacc ggcaccgaat 600

aacatttatc gtaccaccgc agatggtatt ctggcagcac gtgatattga agcgcgtctg 660

aaaaatccgg aaaacctgaa actgctgcaa gaaggtcacg cactggttgg tctgaatccg 720

gaagttggtc atgttcgtat gggttttgaa gatctgccgt atgcatatgc ccgtgttgca 780

cgtgaaggtc gtctgtttca tacccattgg aatagccagc cgctgggtaa ttatgatcag 840

gatctgaata ttggtgtggt ggattgggat agcaccgaag cactgctgta taccctgaaa 900

atggttggtt atcagggcta ttttggcatc gatatcaatc cggaacgcat gccggttatt 960

aaagccattg aaattaacac caaagtgctg cagattatga acgaacgcat tgaacgtctg 1020

ccgcatgatc gtattattga gtgttatttt gaccctgaga atcatcgtgg tgaactggaa 1080

ctgattctgg ccgaaaatca taaa 1104

<210> 17

<211> 1104

<212> DNA

<213> Dictyoglomus thermophilum

<400> 17

atgccgtttg tggatcatcg taatcagaaa attcgtcgca gcaaagaaga actgctgaaa 60

cacatgcaga cctttaaact ggatctgaaa tttagcgttg gcatctggta ttttaccccct 120

ggtggtggtc gttttcatga accgtatgtt gaacagaaag gtattccgga acgtattgaa 180

atggcagcag aaatggcaaa atatggcgtt aaaggtatcg aagcacatta tccggctgaa 240

gtgaatgaag aaaatctgca cctgtataaa cagctggaaa aagaaaccgg tattcgtctg 300

gttgcagttg gtctgagcct gttttatgat aaaatctttg aatttggcag cctgagcaac 360

ccgtatgaaa aatatcgtaa aattgcctat gaacgcctgg tgaatggtct gaaactggtt 420

aaagaagcaa acgccgatat ttgcattatt tggcctggta ttgatggcta tacctatagc 480

tatggtcacc tgtattatca catgtgggat acctttgaag aactggttgc acaggcaatg 540

gatgaagttc cgggtgttca ggttgcaatt gaaccgaaac cgtatgaacc ggcaccgaat 600

aacatttatc gtaccaccgc agatggtatt ctggcagcac gtgatattga agcacgtctg 660

aaaaatccgg aaaacctgaa actgctgcaa gaaggtcacg cactggttgg tctgaatccg 720

gaagttggtc atgttcgtat gggttttgaa gatctgccgt atgcctatgc acgtgttgca 780

cgtgaaggtc gtctgtttca tacccattgg aatagccagc cgctgggtaa ttatgatcag 840

gatctgaata ttggtgtggt ggattgggat agcaccgaag cactgctgta taccctgaaa 900

atggttggtt atcagggcta ttttggcatc gatattaatc cggaacgcat tccggttgtt 960

aaagccattg aaattaacac caaagtgctg cagattatga acgaacgcat tgaacgtctg 1020

ccgcatgatc gtattattga gtgttatttt gaccccgaaa atcatcgtgg tgaactggaa 1080

ctgattctgg cggaaaacca tcgt 1104

<210> 18

<211> 1104

<212> DNA

<213> Dictyoglomus thermophilum

<400> 18

atgccgtttg tggatcatcg taatcagaaa attcgtcgca gcaaagaaga actgctgaaa 60

cacatgcaga cctttaaact ggatctgaaa tttagcgttg gcatctggta ttttaccccct 120

ggtggtggtc gttttcatga accgtatgtt gaacagaaag gtattccgga acgtattgaa 180

atggcagcag aaatggcaaa atatggcgtt aaaggtatcg aagcacatta tccggctgaa 240

gtgaatgaag aaaatctgca cctgtataaa cagctggaaa aagaaaccgg tattcgtctg 300

gttgcagttt ctctgagcct gttttatgat aaaatctttg aatttggcag cctgagcaac 360

ccgtatgaaa aatatcgtaa aattgcctat gaacgcctgg tgaatggtct gaaactggtt 420

aaagaagcaa acgccgatat ttgcattatt tggcctggta ttgatggcta tacctatagc 480

tatggtcacc tgtattatca catgtgggat acctttgaag aactggttgc acaggcaatg 540

gatgaagttc cgggtgttca ggttgcaatt gaaccgaaac cgtatgaacc ggcaccgaat 600

aacatttatc gtaccaccgc agatggtatt ctggcagcac gtgatattga agcacgtctg 660

aaaaatccgg aaaacctgaa actgctgcaa gaaggtcacg cactggttgg tctgaatccg 720

gaagttggtc atgttcgtat gggttttgaa gatctgccgt atgcctatgc acgtgttgca 780

cgtgaaggtc gtctgtttca tacccattgg aatagccagc cgctgggtaa ttatgatcag 840

gatctgaata ttggtgtggt ggattgggat agcaccgaag cactgctgta taccctgaaa 900

atggttggtt atcagggcta ttttggcatc gatattaatc cggaacgcat tccggttgtt 960

aaagccattg aaattaacac caaagtgctg cagattatga acgaacgcat tgaacgtctg 1020

ccgcatgatc gtattattga gtgttatttt gaccccgaaa atcatcgtgg tgaactggaa 1080

ctgattctgg cggaaaacca tcgt 1104

<210> 19

<211> 1104

<212> DNA

<213> Dictyoglomus thermophilum

<400> 19

atgccgtttg tggatcatcg taatcagaaa attcgtcgca gcaaagaaga actgctgaaa 60

cacatgcaga cctttaaact ggatctgaaa tttagcgttg gcatctggta ttttaccccct 120

ggtggtggtc gttttcatga accgtatgtt gaacagaaag gtattccgga acgtattgaa 180

atggcagcag aaatggcaaa atatggcgtt aaaggtatcg aagcacatta tccggctgaa 240

gtgaatgaag aaaatctgca cctgtataaa cagctggaaa aagaaaccgg tattcgtctg 300

gttgcagttg ctctgagcct gttttatgat aaaatctttg aatttggcag cctgagcaac 360

ccgtatgaaa aatatcgtaa aattgcctat gaacgcctgg tgaatggtct gaaactggtt 420

aaagaagcaa acgccgatat ttgcattatt tggcctggta ttgatggcta tacctatagc 480

tatggtcacc tgtattatca catgtgggat acctttgaag aactggttgc acaggcaatg 540

gatgaagttc cgggtgttca ggttgcaatt gaaccgaaac cgtatgaacc ggcaccgaat 600

aacatttatc gtaccaccgc agatggtatt ctggcagcac gtgatattga agcacgtctg 660

aaaaatccgg aaaacctgaa actgctgcaa gaaggtcacg cactggttgg tctgaatccg 720

gaagttggtc atgttcgtat gggttttgaa gatctgccgt atgcctatgc acgtgttgca 780

cgtgaaggtc gtctgtttca tacccattgg aatagccagc cgctgggtaa ttatgatcag 840

gatctgaata ttggtgtggt ggattgggat agcaccgaag cactgctgta taccctgaaa 900

atggttggtt atcagggcta ttttggcatc gatattaatc cggaacgcat tccggttgtt 960

aaagccattg aaattaacac caaagtgctg cagattatga acgaacgcat tgaacgtctg 1020

ccgcatgatc gtattattga gtgttatttt gaccccgaaa atcatcgtgg tgaactggaa 1080

ctgattctgg cggaaaacca tcgt 1104

<210> 20

<211> 1104

<212> DNA

<213> Dictyoglomus thermophilum

<400> 20

atgccgtttg tggatcatcg taatcagaaa attcgtcgca gcaaagaaga actgctgaaa 60

cacatgcaga cctttaaact ggatctgaaa tttagcgttg gcatctggta ttttaccccct 120

ggtggtggtc gttttcatga accgtatgtt gaacagaaag gtattccgga acgtattgaa 180

atggcagcag aaatggcaaa atatggcgtt aaaggtatcg aagcacatta tccggctgaa 240

gtgaatgaag aaaatctgca cctgtataaa cagctggaaa aagaaaccgg tattcgtctg 300

gttgcagttt gcctgagcct gttttatgat aaaatctttg aatttggcag cctgagcaac 360

ccgtatgaaa aatatcgtaa aattgcctat gaacgcctgg tgaatggtct gaaactggtt 420

aaagaagcaa acgccgatat ttgcattatt tggcctggta ttgatggcta tacctatagc 480

tatggtcacc tgtattatca catgtgggat acctttgaag aactggttgc acaggcaatg 540

gatgaagttc cgggtgttca ggttgcaatt gaaccgaaac cgtatgaacc ggcaccgaat 600

aacatttatc gtaccaccgc agatggtatt ctggcagcac gtgatattga agcacgtctg 660

aaaaatccgg aaaacctgaa actgctgcaa gaaggtcacg cactggttgg tctgaatccg 720

```
gaagttggtc atgttcgtat gggttttgaa gatctgccgt atgcctatgc acgtgttgca      780

cgtgaaggtc gtctgtttca tacccattgg aatagccagc cgctgggtaa ttatgatcag      840

gatctgaata ttggtgtggt ggattgggat agcaccgaag cactgctgta taccctgaaa      900

atggttggtt atcagggcta ttttggcatc gatattaatc cggaacgcat tccggttgtt      960

aaagccattg aaattaacac caaagtgctg cagattatga acgaacgcat tgaacgtctg     1020

ccgcatgatc gtattattga gtgttatttt gaccccgaaa atcatcgtgg tgaactggaa     1080

ctgattctgg cggaaaacca tcgt                                           1104
```

REFERENCES CITED IN THE DESCRIPTION

Cited references

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

- EP16175234 [0010] [0031]
- EP2623607A1 [0071]

Non-patent literature cited in the description

- Wood Chemistry**SJOSTROM, E.**Fundamentals and ApplicationsAcademic press19930000292- [0030]
- In vitro mutagenesis protocolsMethods in Molecular BiologyHumana Press20020000vol. 182, [0042]
- Enzyme and Microbial TechnologyElsevier Science Inc19980000vol. V22, 721-723 [0063]

Patentkrav

1. Polypeptid med glucoseisomeraseaktivitet, der omfatter en aminosyresekvens, der er mindst 90 % identisk med aminosyresekvensen ifølge SEQ ID NO: 1 eller SEQ ID NO: 2, hvor polypeptidet omfatter en glycin-, serin-, alanin- eller cysteinrest på en position svarende til position 104 i SEQ ID NO: 1 eller SEQ ID NO: 2.

2. Polypeptid ifølge krav 1, der omfatter en glycinrest på en position svarende til position 104 i SEQ ID NO: 1 eller SEQ ID NO: 2.

3. Polypeptid ifølge krav 1 eller 2, der omfatter en aminosyresekvens, der er mindst 90 % identisk med aminosyren ifølge SEQ ID NO: 1, såsom mindst 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % eller 99 %.

4. Polypeptid ifølge et hvilket som helst af kravene 1 - 3, hvor polypeptidet er et isoleret polypeptid.

5. Sammensætning, der omfatter et polypeptid ifølge et hvilket som helst af kravene 1 - 4.

6. Nukleinsyre, der koder for et polypeptid ifølge et hvilket som helst af kravene 1 - 4.

7. Vektor, der omfatter en nukleinsyre ifølge krav 6.

8. Sammensætning, der omfatter en nukleinsyre ifølge krav 6 eller en vektor ifølge krav 7.

9. Rekombinant værtscelle, der omfatter en nukleinsyre ifølge krav 6, vektor ifølge krav 7 eller sammensætning ifølge krav 8.

10. Rekombinant værtscelle ifølge krav 9 valgt fra gruppen bestående af Escherichia coli, Bacillus, Corynebacterium, Pseudomonas, Pichia pastoris, Saccharomyces cerevisiae, Yarrowia lipolytica, filamentøse svampe, gær og insektceller.

11. Fremgangsmåde til fremstilling af et polypeptid ifølge et hvilket som helst af kravene 1 - 4, hvilken fremgangsmåde omfatter følgende trin:

a. dyrkning af en rekombinant værtscelle ifølge krav 9 eller 10 under forhold, der er egnede til fremstilling af polypeptidet, og

b. genvinding af det opnåede polypeptid og

c. eventuel oprensning af polypeptidet.

12. Anvendelse af et polypeptid ifølge et hvilket som helst af kravene 1 - 4 eller en sammensætning ifølge krav 5 til omdannelse af glucose til fructose, hvor glucosen er afledt af lignocellulosebiomasse.

13. Anvendelse ifølge krav 12, hvor biomassen indeholder xylan eller glucuronoxylan.

14. Anvendelse ifølge krav 12 eller 13, hvor lignocellulosebiomassen er træ, såsom hårdt træ.

15. Fremgangsmåde til øgning af glucoseisomeraseaktivitet for et polypeptid, der omfatter en aminosyresekvens, der er mindst 90 % identisk med aminosyren ifølge SEQ ID NO: 1 eller SEQ ID NO: 2, hvilken fremgangsmåde omfatter trinnet med udskiftning af aminosyren på en position svarende til position 104 i SEQ ID NO: 1 eller SEQ ID NO: 2 med en glycin-, serin-, alanin- eller cysteinrest.

DRAWINGS

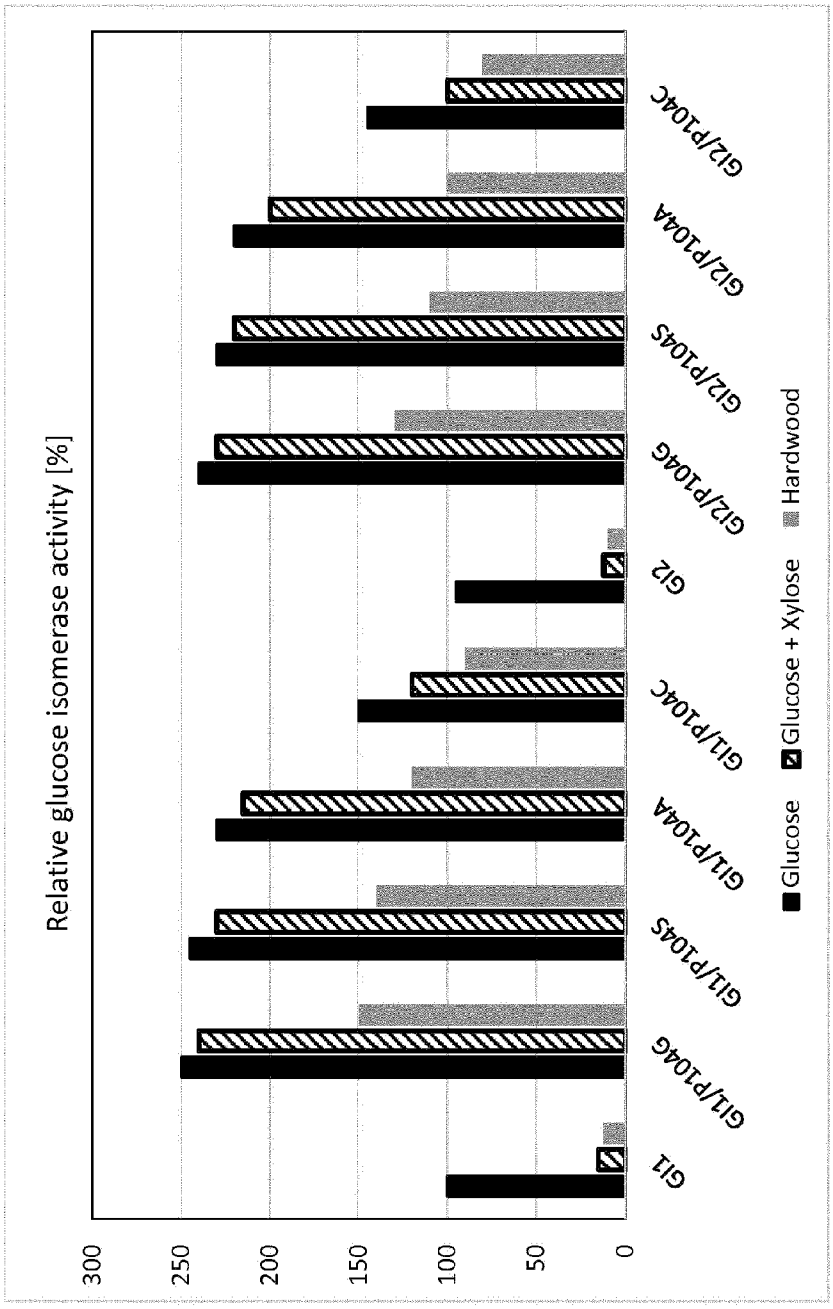

Figure 1