



HU000026367T2

(19) **HU****MAGYARORSZÁG****Szellemi Tulajdon Nemzeti Hivatala**(11) Lajstromszám: **E 026 367**(13) **T2**

EURÓPAI SZABADALOM SZÖVEGÉNEK FORDÍTÁSA

(21) Magyar ügyszám: **E 11 778262**(22) A bejelentés napja: **2011. 05. 04.**(51) Int. Cl.: **A61K 38/16** (2006.01)**C07K 1/00** (2006.01)

(96) Az európai bejelentés bejelentési száma:

EP 20110778262

(86) A nemzetközi (PCT) bejelentési szám:

PCT/US 11/035194

(97) Az európai bejelentés közzétételi adatai:

EP 2566497 A1 **2011. 11. 10.**

(87) A nemzetközi közzétételi szám:

WO 11140219

(97) Az európai szabadalom megadásának meghirdetési adatai:

EP 2566497 B1 **2015. 07. 29.**

(30) Elsőbbségi adatok:

331245 P **2010. 05. 04.** **US**

(73) Jogosult(ak):

Codexis, Inc., Redwood City, CA 94063 (US)

(72) Feltaláló(k):

CROWE, Michael, A., Singapore 117525 (SG)
ALVIZO, Oscar, Redwood City California 94063 (US)
BEHROUZIAN, Behnaz, Redwood City California 94063 (US)
BONG, Yong Koy, Singapore 117525 (SG)
COLLIER, Steven, J., Singapore 248474 (SG)
GOHEL, Anupam, Singapore 117525 (SG)
MAVINAHALLI, Jagadeesh, Singapore 117525 (SG)
MODUKURU, Naga, Singapore 117525 (SG)
MUNDORFF, Emily, Belmont California 94002 (US)
SMITH, Derek, Singapore 117525 (SG)
SONG, Shiwei, Singapore 117525 (SG)
YEO, Wan Lin, Singapore 117525 (SG)

(74) Képviselő:

dr. Kompagne Hajnalka, DANUBIA Szabadalmi és Jogi Iroda Kft., Budapest

(54)

Ezetimib-szintézis biokatalizátorai

Az európai szabadalom ellen, megadásának az Európai Szabadalmi Közlönyben való meghirdetésétől számított kilenc hónapon belül, felszólalást lehet benyújtani az Európai Szabadalmi Hivatalnál. (Európai Szabadalmi Egyezmény 99. cikk(1))

A fordítást a szabadalmas az 1995. évi XXXIII. törvény 84/H. §-a szerint nyújtotta be. A fordítás tartalmi helyességét a Szellemi Tulajdon Nemzeti Hivatala nem vizsgálta.



(11) **EP 2 566 497 B1**

(12) **EUROPEAN PATENT SPECIFICATION**

(45) Date of publication and mention
of the grant of the patent:
29.07.2015 Bulletin 2015/31

(21) Application number: **11778262.3**

(22) Date of filing: **04.05.2011**

(51) Int Cl.:
A61K 38/16^(2006.01) C07K 1/00^(2006.01)

(86) International application number:
PCT/US2011/035194

(87) International publication number:
WO 2011/140219 (10.11.2011 Gazette 2011/45)

(54) **BIOCATALYSTS FOR EZETIMIBE SYNTHESIS**

BIOKATALYSATOREN ZUR EZETIMIB-SYNTHESE

BIOCATALYSEURS POUR LA SYNTHÈSE D'ÉZÉTIMIBE

(84) Designated Contracting States:
**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**

(30) Priority: **04.05.2010 US 331245 P**

(43) Date of publication of application:
13.03.2013 Bulletin 2013/11

(73) Proprietor: **Codexis, Inc.**
Redwood City, CA 94063 (US)

(72) Inventors:

- **CROWE, Michael, A.**
Singapore 117525 (SG)
- **ALVIZO, Oscar**
Redwood City
California 94063 (US)
- **BEHROUZIAN, Behnaz**
Redwood City
California 94063 (US)
- **BONG, Yong Koy**
Singapore 117525 (SG)
- **COLLIER, Steven, J.**
Singapore 248474 (SG)
- **GOHEL, Anupam**
Singapore 117525 (SG)
- **MAVINAHALLI, Jagadeesh**
Singapore 117525 (SG)
- **MODUKURU, Naga**
Singapore 117525 (SG)
- **MUNDORFF, Emily**
Belmont
California 94002 (US)

- **SMITH, Derek**
Singapore 117525 (SG)
- **SONG, Shiwei**
Singapore 117525 (SG)
- **YEO, Wan Lin**
Singapore 117525 (SG)

(74) Representative: **Chapman, Desmond Mark**
Carpmaels & Ransford LLP
One Southampton Row
London WC1B 5HA (GB)

(56) References cited:

WO-A2-2009/036404 WO-A2-2010/025085
WO-A2-2010/025238 US-A1- 2008 318 295
US-A1- 2010 062 499

- **DATABASE Geneseq [Online] 29 April 2010**
(2010-04-29), "Lactobacillus kefir ketoreductase
(KRED) variant protein SEQ ID 96.",
XP002714446, retrieved from EBI accession no.
GSP:AXW52051 Database accession no.
AXW52051
- **DATABASE Geneseq [Online] 29 April 2010**
(2010-04-29), "Lactobacillus kefir ketoreductase
(KRED) variant protein SEQ ID 124.",
XP002714447, retrieved from EBI accession no.
GSP:AXW52079 Database accession no.
AXW52079
- **DATABASE Geneseq [Online] 14 May 2009**
(2009-05-14), "Lactobacillus kefir mutant
ketoreductase protein sequence, SEQ ID 6.",
XP002714448, retrieved from EBI accession no.
GSP:AWI71982 Database accession no.
AWI71982

Note: Within nine months of the publication of the mention of the grant of the European patent in the European Patent Bulletin, any person may give notice to the European Patent Office of opposition to that patent, in accordance with the Implementing Regulations. Notice of opposition shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

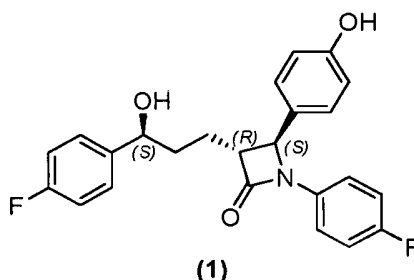
Description

1. TECHNICAL FIELD

[0001] The present disclosure relates to biocatalysts and methods of using the biocatalysts.

2. BACKGROUND

[0002] The present disclosure relates to improved biocatalysts and improved biocatalytic processes for the preparation of the active pharmaceutical ingredient, (1-(4-fluorophenyl)-3(R)-[3-(4-fluorophenyl)-3(S)-hydroxypropyl]-4(S)-(4-hydroxyphenyl)-2-azetidinone) (shown below as compound (1)) and derivatives and analogs thereof.



[0003] Compound (1) is commonly known as Ezetimibe and is the active ingredient in ZETIA®, manufactured by Merck/Schering-Plough Pharmaceuticals. Ezetimibe has been approved by the United States Food and Drug Administration for use in patients with high cholesterol to reduce LDL cholesterol and total cholesterol (see e.g., US 6,207,822). Ezetimibe lowers high levels of blood cholesterol by selectively inhibiting the intestinal absorption of cholesterol and related phytosterols. Ezetimibe is commercially available in combination with simvastatin in the VYTORIN™ formulation from MSP Pharmaceuticals, Inc.

[0004] Numerous compounds that are analogs of Ezetimibe and being developed as possible therapeutics for lowering cholesterol are also known in the art (see e.g., PCT publications WO2006/17257A2, WO 2008/085300A1, and WO2008/039829A2).

[0005] Synthetic processes for the production of Ezetimibe and Ezetimibe derivatives have been previously disclosed. A variety of publications have disclosed chemical synthesis using a late reduction scheme that delays the reduction of the alcohol to the carbonyl to the last step of the reaction: US 5,886,171, US 5,738,321, WO 2005/0066120, WO 2007/030721, WO 2007/120824, WO 2007/119106, WO 2007/072088, WO 2007/030721, and WO 2007/120824.

[0006] US 6,133,001 and WO 2000/060107 disclose using certain microorganisms (e.g., *Rhodococcus fascians* ATCC No. 202210 or *Geotrichum candidum* ATCC No. 74487) to carry out the stereoselective reduction of 1-(4-fluorophenyl)-3(R)-[3-oxo-3-(4-fluorophenyl)propyl]-4(S)-(4-hydroxyphenyl)-2-azetidinone to 1-(4-fluorophenyl)-3(R)-[3(S)-hydroxy-3-(4-fluorophenyl)-propyl]-4(S)-(4-hydroxyphenyl)-2-azetidinone. This is a microbial process, however, carried out under whole cell fermentation conditions.

[0007] WO 2008/151324A1 discloses using certain commercially available ketoreductase biocatalysts to prepare Ezetimibe and protected Ezetimibe analogs from the corresponding precursor ketone compounds. The biocatalysts and processes disclosed therein, however, use low substrate loadings (25 g/L or less), a GDH/glucose cofactor regeneration system, and result in low percentage conversion of substrate to the Ezetimibe product (~65% yield).

[0008] US20100062499A1 discloses engineered ketoreductase enzymes, and methods of using the engineered ketoreductase enzymes to convert the diketone compound, 5-((4S)-2-oxo-4-phenyl(1,3-oxazolidin-3-yl))-1-(4-fluorophenyl)pentane-1,5-dione, to the chiral alcohol, (4S)-3-[(5S)-5-(4-fluorophenyl)-5-hydroxypentanoyl]-4-phenyl-1,3-oxazolidin-2-one. This chiral alcohol made biocatalytically is an early stage intermediate that can be used in a process for making Ezetimibe.

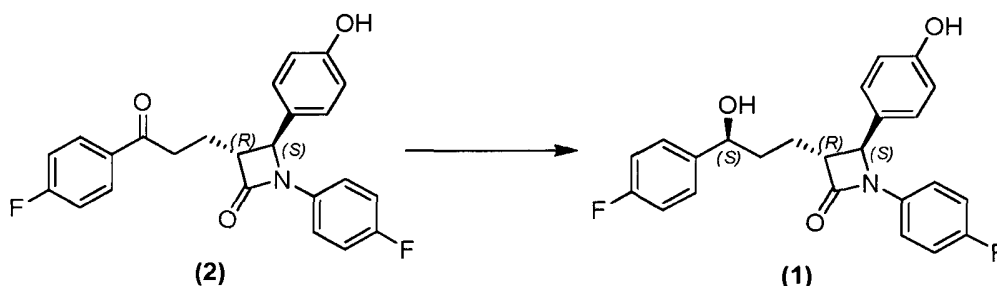
[0009] WO2010/025085, WO2009/036404, and EBI accessions AXW52051, AXW52079 and AWI71982 disclose *Lactobacillus* KRED sequences.

[0010] It is desirable to have improved biocatalysts and a biocatalytic process having increased efficiency for use in a late stage biocatalytic reduction scheme for preparing Ezetimibe in high diastereomeric excess (> 98% d.e.). Particularly desirable would be engineered biocatalysts capable of increased activity in large scale processes having high substrate loadings (e.g., > 50 g/L), high percent conversion (e.g., > 90% in 24 h), without the need for an additional cofactor regenerating enzyme, and capable of yielding Ezetimibe as product in high purity and diastereomeric excess.

3. SUMMARY

[0011] The present disclosure provides non-naturally occurring polypeptides having ketoreductase activity, polynucleotides encoding the polypeptides, methods of the making the polypeptides, and methods of using the polypeptides for the biocatalytic conversion of the keto-phenol substrate, 1-(4-fluorophenyl)-3(R)-[3-(4-fluorophenyl)-3-oxopropyl]-4(S)-(4-hydroxyphenyl)-2-azetidinone (compound (2) below) to the chiral (S)-alcohol product, 1-(4-fluorophenyl)-3(R)-[3-(4-fluorophenyl)-3(S)-hydroxypropyl]-4(S)-(4-hydroxyphenyl)-2-azetidinone (*i.e.*, compound (1) commonly referred to as Ezetimibe) as shown in Scheme 1.

Scheme 1



[0012] While naturally occurring ketoreductase polypeptides do not efficiently convert compound (2) to compound (1), the non-naturally occurring, engineered, ketoreductase polypeptides of the present disclosure are capable of carrying out this conversion with improved properties including, high diastereomeric excess (*e.g.*, at least about 99% d.e.), increased activity (*e.g.*, at least about 10-fold increased activity relative to the reference polypeptide SEQ ID NO:2), high percent conversion (*e.g.*, at least about 90% conversion in 24 h), in the presence of high substrate loadings (*e.g.*, at least about 50 g/L compound (2)), and without any cofactor regenerating enzyme other than the engineered ketoreductase polypeptide.

[0013] The non-naturally occurring polypeptides of the present disclosure capable of converting compound (2) to compound (1) with at least 2-fold, at least 10-fold, at least 25-fold, at least 40-fold, or at least 60-fold increased activity relative to the activity of the reference polypeptide of SEQ ID NO: 2, are synthetic variants of the naturally occurring ketoreductase of *Lactobacillus kefir*, and comprise amino acid sequences that have one or more residue differences as compared to the reference sequence of the synthetic variant ketoreductase polypeptide of SEQ ID NO:2. The residue differences occur at residue positions that affect functional properties of the enzyme including activity (*e.g.*, percent conversion of substrate to product), stereoselectivity, substrate and/or product binding (*e.g.*, resistance to substrate and/or product inhibition), thermostability, solvent stability, expression, or various combinations thereof. Accordingly, in some embodiments, the polypeptides of the disclosure can have one or more residue differences as compared to SEQ ID NO:2 at the following residue positions: **X21, X25, X40, X64, X93, X94, X95, X96, X99, X108, X117, X127, X147, X148, X150, X152, X153, X155, X190, X195, X196, X201, X202, X203, X204, X205, X206, X207, X211, X221, X223, and X226.** Amino acid residues that can be present at these positions are described in detail in the descriptions herein.

[0014] In some embodiments, the present disclosure provides a non-naturally occurring polypeptide of capable of converting compound (2) to compound (1) comprising an amino acid sequence selected from any one of SEQ ID NO: 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, or 168.

[0015] In some embodiments, the present disclosure provides a non-naturally occurring polypeptide capable of converting compound (2) to compound (1) with at least 2-fold, at least 10-fold, at least 25-fold, at least 40-fold, or at least 60-fold increased activity relative to the activity of the polypeptide of SEQ ID NO: 2, and comprises an amino acid sequence having at least 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identity to a reference amino acid sequence selected from any one of SEQ ID NO: 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, or 168.

[0016] In some embodiments, the present disclosure provides a non-naturally occurring polypeptide capable of converting compound (2) to compound (1) with at least 2-fold, at least 10-fold, at least 25-fold, at least 40-fold, or at least

60-fold increased activity relative to the activity of the polypeptide of SEQ ID NO: 2, comprises an amino acid sequence having at least 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identity to SEQ ID NO: 2, and further comprises a set of amino acid residue differences as compared to SEQ ID NO: 2 of any one of SEQ ID NO: 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, or 168. In some embodiments, in addition to the set of amino acid residue differences of any one of the non-naturally occurring polypeptides of SEQ ID NO: 4 through SEQ ID NO: 168, the sequence of the non-naturally occurring polypeptide can further comprise 1-2, 1-3, 1-4, 1-5, 1-6, 1-7, 1-8, 1-9, 1-10, 1-11, 1-12, 1-14, 1-15, 1-16, 1-18, 1-20, 1-22, 1-24, 1-26, 1-30, 1-35, 1-40 residue differences at other amino acid residue positions as compared to the SEQ ID NO: 2. In some embodiments, the residue differences can comprise conservative substitutions and non-conservative substitutions as compared to SEQ ID NO: 2.

[0017] In some embodiments, the present disclosure provides a non-naturally occurring polypeptide capable of converting compound (2) to compound (1) with at least 2-fold, at least 10-fold, at least 25-fold, at least 40-fold, or at least 60-fold increased activity relative to the activity of the polypeptide of SEQ ID NO: 2, which comprises an amino acid sequence having at least 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identity to SEQ ID NO: 2 and at least the following features: residue at position corresponding to **X40** is R; residue at position corresponding to **X153** is I, or L; residue at position corresponding to **X190** is A or P; residue at position corresponding to **X196** is T; residue at position corresponding to **X199** is F, or W; and residue at position corresponding to **X206** is I. In some embodiments, the amino acid sequence further comprises at least one feature or group of features selected from: (a) residue at position **X93** is A and residue at position **X94** is T; (b) residue at position **X93** is A and residue at position **X94** is S; (c) residue at position **X93** is A and residue at position **X94** is S; (d) residue at position **X93** is I and residue at position **X94** is S; (e) residue at position **X203** is G; (f) residue at position **X202** is G and residue at position **X203** is G; or (f) residue at position **X201** is A, residue at position **X202** is G, and residue at position **X203** is G.

[0018] In some embodiments, any of the non-naturally occurring polypeptides of the present disclosure capable of converting compound (2) to compound (1) with at least 2-fold, at least 10-fold, at least 25-fold, at least 40-fold, or at least 60-fold increased activity relative to the activity of the polypeptide of SEQ ID NO: 2 and an amino acid sequence having at least 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identity to SEQ ID NO: 2, can further one or more features selected from: residue at position corresponding to **X21** is R or F; residue at position corresponding to **X25** is R, T, or N; residue at position corresponding to **X40** is R; residue at position corresponding to **X64** is V; residue at position corresponding to **X93** is A; residue at position corresponding to **X94** is T, S, or P; residue at position corresponding to **X95** is V, or M; residue at position corresponding to **X96** is V, G, A, N, S, P, or T; residue at position corresponding to **X99** is L; residue at position corresponding to **X108** is H; residue at position corresponding to **X117** is A, or G; residue at position corresponding to **X127** is K, or Q; residue at position corresponding to **X147** is M, or I; residue at position corresponding to **X148** is I; residue at position corresponding to **X150** is H, or A; residue at position corresponding to **X152** is N, or F; residue at position corresponding to **X153** is I, or L; residue at position corresponding to **X155** is C; residue at position corresponding to **X190** is A; residue at position corresponding to **X195** is M; residue at position corresponding to **X196** is T, A, S, C, or N; residue at position corresponding to **X199** is F, or W; residue at position corresponding to **X201** is I, L, or A; residue at position corresponding to **X202** is L, N, V, or G; residue at position corresponding to **X203** is G; residue at position corresponding to **X204** is V, or A; residue at position corresponding to **X205** is V; residue at position corresponding to **X206** is I; residue at position corresponding to **X207** is T, C, I, or N; residue at position corresponding to **X211** is K; residue at position corresponding to **X221** is D; residue at position corresponding to **X223** is I; residue at position corresponding to **X226** is V.

[0019] In some embodiments, the non-naturally occurring polypeptides of the present disclosure can comprise a sequence having various combinations of the residue differences as compared to SEQ ID NO: 2 disclosed herein at residue positions affecting enzymatic activity, thermostability, solvent stability, and cofactor binding.

[0020] In some embodiments, the non-naturally occurring polypeptides of the present disclosure can comprise a sequence having one or more amino acid residue differences as compared to SEQ ID NO: 2 at residue positions affecting activity for conversion of compound (2) to compound (1), are selected from the following position: X21; X25; X64; X93; X94; X95; X96; X99; X108; X117; X127; X147; X148; X150; X152; X153; X155; X163; X190; X195; X196; X199; X201; X202; X203; X204; X205; X206; X207; X211; X221; X223; and X226. In some embodiments, specific amino acid differences at residue positions resulting in increased activity for conversion of compound (2) to compound (1) relative to the reference polypeptide of SEQ ID NO: 2 can be selected from the following substitutions: L21FR; D25NRT; A64V; I93AT; A94PST; L95MV; Q96ANGPSTV; V99L; R108DHK; S117AG; R127KQ; L147IM; V148I; D150H; M152NF; V153IL; A155C; V163I; C190A; L195M; V196ACNST; D199FW; G201AIL; A202GLNV; E203G; E204AV; M205V; M206I; S207TCIN; R211K; N221D; V223I; and I226V.

[0021] In some embodiments, the non-naturally occurring polypeptides of the present disclosure can comprise an amino acid sequence having one or more residue differences as compared to SEQ ID NO: 2 at residue positions affecting

thermostability, which positions include the following: X21; X93; X94; X117; X127; X147; X195; and X199. In some embodiments, specific amino acid differences at residue positions resulting in increased thermostability relative to the reference polypeptide of SEQ ID NO: 2 can be selected from the following substitutions: L21F; A93T; S94A; S117GA; R127K; L147I; L195M; and D199W.

[0022] In some embodiments, the non-naturally occurring polypeptides of the present disclosure can comprise an amino acid sequence having residue differences as compared to SEQ ID NO: 2 at residue positions affecting solvent stability, which positions include X25; X147; and X221. In some embodiments, specific amino acid differences at residue positions resulting in increased solvent stability relative to the reference polypeptide of SEQ ID NO: 2 can be selected from the following substitutions: D25R; L147M; and N221D.

[0023] In some embodiments, the non-naturally occurring polypeptides of the present disclosure can comprise an amino acid sequence having residue differences as compared to SEQ ID NO: 2 at residue positions affecting cofactor binding, which positions include X40. In some embodiments, specific amino acid differences at residue positions affecting cofactor binding can be selected from the following substitutions: H40R.

[0024] In addition to the residue position specified above, various other residue differences relative to SEQ ID NO: 2 can be present at other residue positions in the ketoreductase polypeptides disclosed herein. These can be conservative or non-conservative differences, including conservative substitutions and non-conservative substitutions. Guidance on the choice of amino acid residues at the specified positions is provided in the detailed description.

[0025] In some embodiments, the present disclosure provides polynucleotides encoding the non-naturally occurring polypeptides capable of converting compound (2) to compound (1), as well as expression vectors comprising the polynucleotides, and host cells capable of expressing the polynucleotides encoding the non-naturally occurring polypeptides. Accordingly, in some embodiments, the present disclosure also provides methods of manufacturing the non-naturally occurring polypeptides capable of converting compound (2) to compound (1), wherein the methods comprise culturing a host cell capable of expressing a polynucleotide encoding the non-naturally occurring polypeptide and isolating the polypeptide from the host cell.

[0026] In some embodiments, any of the non-naturally occurring polypeptides of the present disclosure can be used in improved processes for carrying out the conversion of compound (2) to compound (1) due to their improved enzymatic properties including, high diastereomeric excess (*e.g.*, at least about 99% d.e.), increased activity (*e.g.*, at least about 10-fold increased activity relative to SEQ ID NO: 2), high percent conversion (*e.g.*, at least about 90% conversion in 24 h), in the presence of high substrate loadings (*e.g.*, at least about 50 g/L compound (2)), and without any cofactor regenerating enzyme other than the non-naturally occurring ketoreductase polypeptide. Accordingly, in some embodiments, the present disclosure provides methods using the non-naturally occurring polypeptides for preparing compound (1) in diastereomeric excess, wherein the methods comprise: contacting compound (2) with an non-naturally occurring or engineered polypeptide of the present disclosure (*e.g.*, as described in Table 2 and elsewhere herein) in the presence of NADPH or NADH cofactor under suitable reaction conditions. Suitable reaction conditions for the conversion of compound (2) to compound (1) using the engineered polypeptides of the present disclosure are described in greater detail below, including but not limited to ranges of pH, temperature, buffer, solvent system, substrate loading, polypeptide loading, cofactor loading, atmosphere, and reaction time.

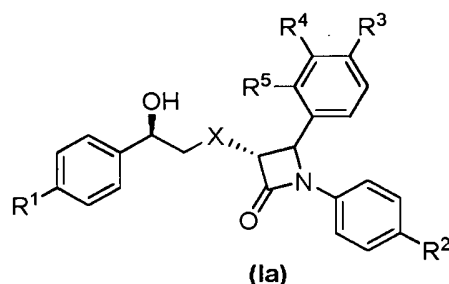
[0027] In some embodiments, the improved enzymatic activity of the engineered polypeptides in the conversion of compound (2) to compound (1) provides for methods wherein a higher percentage conversion can be achieved with a lower concentration of polypeptide. The use of lower concentration of the engineered polypeptide in a method comprising a conversion of compound (2) to compound (1) also reduces the amount of residual protein that may need to be removed in subsequent steps for purification of compound (1). Accordingly, in some embodiments, the methods for preparing compound (1) of the present disclosure can be carried out wherein the suitable reaction conditions can comprise *e.g.*, a compound (2) substrate loading of at least about 20 g/L, about 40 g/L, about 50 g/L, about 75 g/L, about 100 g/L, about 200 g/L, about 250 g/L, about 300 g/L, or about 400 g/L; and/or a non-naturally occurring polypeptide concentration of about 0.1-3.0 g/L, about 0.5-2.75 g/L, about 1.0-2.5 g/L, about 1.5-2.5 g/L, about 3 g/L, about 2 g/L, about 1.5 g/L, about 1.0 g/L, about 0.75 g/L, or even lower concentration.

[0028] In some embodiments, the present disclosure also provides methods for preparing compound (1) or an analog of compound (1), wherein the methods comprise contacting compound (2) or an analog of compound (2) with a non-naturally occurring or engineered polypeptide in the presence of NADPH or NADH cofactor under suitable reaction conditions and further comprises chemical steps of product work-up, extraction, isolation, purification, and/or crystallization of compound (1), each of which can be carried out under a range of conditions disclosed herein.

[0029] In some embodiments, the methods for preparing compound (1) using a non-naturally occurring polypeptide of the present disclosure further comprise a cofactor recycling system capable of converting NADP⁺ to NADPH, or NAD⁺ to NADH. The cofactor recycling system can comprise a dehydrogenase enzyme (*e.g.*, glucose dehydrogenase, glucose-phosphate dehydrogenase, formate dehydrogenase, or a ketoreductase/alcohol dehydrogenase) and a corresponding substrate (*e.g.*, glucose, glucose-6-phosphate, formate, or secondary alcohol). In some embodiments, the co-factor recycling system comprises a ketoreductase polypeptide and a secondary alcohol, preferably isopropanol. In some

embodiments of the methods of the present disclosure, the non-naturally occurring polypeptide capable of converting compound (2) to compound (1) is also capable of converting a secondary alcohol (e.g., isopropanol) to its corresponding secondary ketone (e.g., acetone), and the method of preparing compound (1) further comprises a co-factor recycling system comprising the non-naturally occurring polypeptide and a secondary alcohol.

[0030] In some embodiments, an analog of compound (1) can be prepared in diastereomeric excess from an analog of compound (2) using the above described methods. In some embodiments, the analog of compound (1) prepared using the methods comprises a compound of Formula **1a**:



wherein,

X is C or S;

R¹ is selected from -H, -F, -Cl, -Br, or -I;

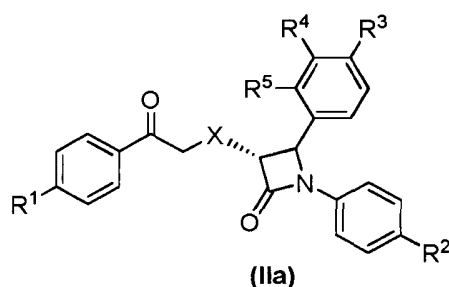
R² is selected from -H, -F, -Cl, -Br, -I, -CN, -OH (optionally protected with a hydroxyl protecting group), -CH₂NH₂ (optionally protected with a nitrogen protecting group), and any one of the following optionally substituted groups: alkyl, alkoxy, alkenyl, alkenoxy, alkynyl, alkynoxy, cycloalkyl, aryl, heteroaryl, or heterocycle;

R³ is selected from -H, -F, -Cl, -Br, -I, -CN, -OH (optionally protected with a hydroxyl protecting group), -CH₂NH₂ (optionally protected with a nitrogen protecting group), and any one of the following optionally substituted groups: alkyl, alkoxy, alkenyl, alkenoxy, alkynyl, alkynoxy, cycloalkyl, aryl, heteroaryl, or heterocycle;

R⁴ is selected from -H, -F, -Cl, -Br, -I, -CN, -OH (optionally protected with a hydroxyl protecting group); and

R⁵ is selected from -H, -F, -Cl, -Br, -I, -CN, -OH (optionally protected with a hydroxyl protecting group), -CH₂NH₂ (optionally protected with a nitrogen protecting group), and any one of the following optionally substituted groups: alkyl, alkoxy, alkenyl, alkenoxy, alkynyl, alkynoxy, cycloalkyl, aryl, heteroaryl, or heterocycle.

[0031] Accordingly, in some embodiments the present disclosure provides a method of preparing a compound of Formula **1a** in diastereomeric excess comprising: contacting a compound of Formula **1a**



wherein, X, R¹, R², R³, R⁴, and R⁵, are defined as above for Formula **1a**, with an engineered polypeptide of the present disclosure (e.g., as described in Table 2 and elsewhere herein) in the presence of NADPH or NADH cofactor under suitable reaction conditions.

4. DETAILED DESCRIPTION

[0032] The present disclosure provides highly stereoselective and efficient biocatalysts capable of mediating the conversion of 1-(4-fluorophenyl)-3(R)-[3-(4-fluorophenyl)-3-oxopropyl]-4(S)-(4-hydroxyphenyl)-2-azetidinone to 1-(4-fluorophenyl)-3(R)-[3-(4-fluorophenyl)-3(S)-hydroxypropyl]-4(S)-(4-hydroxyphenyl)-2-azetidinone in diastereomeric excess. The biocatalysts described herein have been designed by changing the amino acid sequence of a naturally occurring

ketoreductase to form polypeptides with the desired enzymatic properties, e.g., enzyme activity, stereoselectivity, by product formation, thermostability, and expression. The detailed description that follow describes the polypeptides and processes for carrying out the conversion of 1-(4-fluorophenyl)-3(R)-[3-(4-fluorophenyl)-3-oxopropyl]-4(S)-(4-hydroxyphenyl)-2-azetidinone to 1-(4-fluorophenyl)-3(R)-[3-(4-fluorophenyl)-3(S)-hydroxypropyl]-4(S)-(4-hydroxyphenyl)-2-azetidinone in diastereomeric excess.

[0033] For the descriptions herein and the appended claims, the singular forms "a", "an" and "the" include plural referents unless the context clearly indicates otherwise. Thus, for example, reference to "a polypeptide" includes more than one polypeptide, and reference to "a compound" refers to more than one compound.

[0034] Also, the use of "or" means "and/or" unless stated otherwise. Similarly, "comprise," "comprises," "comprising" "include," "includes," and "including" are interchangeable and not intended to be limiting.

[0035] It is to be further understood that where descriptions of various embodiments use the term "comprising," those skilled in the art would understand that in some specific instances, an embodiment can be alternatively described using language "consisting essentially of" or "consisting of."

[0036] It is to be understood that both the foregoing general description, including the drawings, and the following detailed description are exemplary and explanatory only and are not restrictive of this disclosure.

4.2 Definitions

[0037] The technical and scientific terms used in the descriptions herein will have the meanings commonly understood by one of ordinary skill in the art, unless specifically defined otherwise. Accordingly, the following terms are intended to have the following meanings.

[0038] "Protein", "polypeptide," and "peptide" are used interchangeably herein to denote a polymer of at least two amino acids covalently linked by an amide bond, regardless of length or post-translational modification (e.g., glycosylation, phosphorylation, lipidation, myristilation, ubiquitination, etc.). Included within this definition are D- and L-amino acids, and mixtures of D- and L-amino acids.

[0039] "Coding sequence" refers to that portion of a nucleic acid (e.g., a gene) that encodes an amino acid sequence of a protein.

[0040] "Naturally occurring" or "wild-type" refers to the form found in nature. For example, a naturally occurring or wild-type polypeptide or polynucleotide sequence is a sequence present in an organism that can be isolated from a source in nature and which has not been intentionally modified by human manipulation.

[0041] "Non-naturally occurring" or "engineered" or "recombinant" when used in the present disclosure with reference to, e.g., a cell, nucleic acid, or polypeptide, refers to a material, or a material corresponding to the natural or native form of the material, that has been modified in a manner that would not otherwise exist in nature, or is identical thereto but produced or derived from synthetic materials and/or by manipulation using recombinant techniques. Non-limiting examples include, among others, recombinant cells expressing genes that are not found within the native (non-recombinant) form of the cell or express native genes that are otherwise expressed at a different level.

[0042] "Percentage of sequence identity," "percent identity," and "percent identical" are used herein to refer to comparisons between polynucleotide sequences or polypeptide sequences, and are determined by comparing two optimally aligned sequences over a comparison window, wherein the portion of the polynucleotide or polypeptide sequence in the comparison window may comprise additions or deletions (*i.e.*, gaps) as compared to the reference sequence for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which either the identical nucleic acid base or amino acid residue occurs in both sequences or a nucleic acid base or amino acid residue is aligned with a gap to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison and multiplying the result by 100 to yield the percentage of sequence identity. Determination of optimal alignment and percent sequence identity is performed using the BLAST and BLAST 2.0 algorithms (see e.g., Altschul et al., 1990, J. Mol. Biol. 215: 403-410 and Altschul et al., 1977, Nucleic Acids Res. 3389-3402). Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information website.

[0043] Briefly, the BLAST analyses involve first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as, the neighborhood word score threshold (Altschul et al, supra). These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are then extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always >0) and N (penalty score for mismatching residues; always <0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring

residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) of 10, M=5, N=-4, and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength (W) of 3, an expectation (E) of 10, and the BLOSUM62 scoring matrix (see Henikoff and Henikoff, 1989, Proc Natl Acad Sci USA 89:10915).

[0044] Numerous other algorithms are available that function similarly to BLAST in providing percent identity for two sequences. Optimal alignment of sequences for comparison can be conducted, e.g., by the local homology algorithm of Smith and Waterman, 1981, Adv. Appl. Math. 2:482, by the homology alignment algorithm of Needleman and Wunsch, 1970, J. Mol. Biol. 48:443, by the search for similarity method of Pearson and Lipman, 1988, Proc. Natl. Acad. Sci. USA 85:2444, by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the GCG Wisconsin Software Package), or by visual inspection (see generally, Current Protocols in Molecular Biology, F. M. Ausubel et al., eds., Current Protocols, a joint venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc., (1995 Supplement) (Ausubel)). Additionally, determination of sequence alignment and percent sequence identity can employ the BESTFIT or GAP programs in the GCG Wisconsin Software package (Accelrys, Madison WI), using default parameters provided.

[0045] "Reference sequence" refers to a defined sequence to which another sequence is compared. A reference sequence may be a subset of a larger sequence, for example, a segment of a full-length gene or polypeptide sequence. Generally, a reference sequence is at least 20 nucleotide or amino acid residues in length, at least 25 residues in length, at least 50 residues in length, or the full length of the nucleic acid or polypeptide. Since two polynucleotides or polypeptides may each (1) comprise a sequence (*i.e.*, a portion of the complete sequence) that is similar between the two sequences, and (2) may further comprise a sequence that is divergent between the two sequences, sequence comparisons between two (or more) polynucleotides or polypeptide are typically performed by comparing sequences of the two polynucleotides over a comparison window to identify and compare local regions of sequence similarity. The term "reference sequence" is not intended to be limited to wild-type sequences, and can include engineered or altered sequences. For example, in some embodiments, a "reference sequence" can be a previously engineered or altered amino acid sequence.

[0046] "Comparison window" refers to a conceptual segment of at least about 20 contiguous nucleotide positions or amino acids residues wherein a sequence may be compared to a reference sequence of at least 20 contiguous nucleotides or amino acids and wherein the portion of the sequence in the comparison window may comprise additions or deletions (*i.e.*, gaps) of 20 percent or less as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. The comparison window can be longer than 20 contiguous residues, and includes, optionally 30, 40, 50, 100, or longer windows.

[0047] "Corresponding to", "reference to" or "relative to" when used in the context of the numbering of a given amino acid or polynucleotide sequence refers to the numbering of the residues of a specified reference sequence when the given amino acid or polynucleotide sequence is compared to the reference sequence. In other words, the residue number or residue position of a given polymer is designated with respect to the reference sequence rather than by the actual numerical position of the residue within the given amino acid or polynucleotide sequence. For example, a given amino acid sequence, such as that of an engineered ketoreductase, can be aligned to a reference sequence by introducing gaps to optimize residue matches between the two sequences. In these cases, although the gaps are present, the numbering of the residue in the given amino acid or polynucleotide sequence is made with respect to the reference sequence to which it has been aligned.

[0048] "Stereoselectivity" refers to the preferential formation in a chemical or enzymatic reaction of one stereoisomer over another. Stereoselectivity can be partial, where the formation of one stereoisomer is favored over the other, or it may be complete where only one stereoisomer is formed. When the stereoisomers are enantiomers, the stereoselectivity is referred to as stereoselectivity, the fraction (typically reported as a percentage) of one enantiomer in the sum of both. It is commonly alternatively reported in the art (typically as a percentage) as the enantiomeric excess (*e.e.*) calculated therefrom according to the formula $[\text{major enantiomer} - \text{minor enantiomer}] / [\text{major enantiomer} + \text{minor enantiomer}]$. Where the stereoisomers are diastereoisomers, the stereoselectivity is referred to as diastereoselectivity, the fraction (typically reported as a percentage) of one diastereomer in a mixture of two diastereomers, commonly alternatively reported as the diastereomeric excess (*d.e.*). Enantiomeric excess and diastereomeric excess are types of stereomeric excess.

[0049] "Highly stereoselective" refers to a chemical or enzymatic reaction that is capable of converting a substrate (*e.g.*, compound **(2)**, 1-(4-fluorophenyl)-3(R)-[3-(4-fluorophenyl)-3-oxopropyl]-4(S)-(4-hydroxyphenyl)-2-azetidinone) to its corresponding product (*e.g.*, compound **(1)**, 1-(4-fluorophenyl)-3(R)-[3-(4-fluorophenyl)-3(S)-hydroxypropyl]-4(S)-(4-hydroxyphenyl)-2-azetidinone) with at least about 85% stereoisomeric excess.

[0050] "Increased enzymatic activity" or "increased activity" refers to an improved property of an engineered enzyme, which can be represented by an increase in specific activity (*e.g.*, product produced/time/weight protein) or an increase in percent conversion of the substrate to the product (*e.g.*, percent conversion of starting amount of substrate to product in a specified time period using a specified amount of ketoreductase) as compared to a reference enzyme. Exemplary methods to determine enzyme activity are provided in the Examples. Any property relating to enzyme activity may be

affected, including the classical enzyme properties of K_m , V_{max} or k_{cat} , changes of which can lead to increased enzymatic activity. The ketoreductase activity can be measured by any one of standard assays used for measuring ketoreductases, such as change in substrate or product concentration, or change in concentration of the cofactor (in absence of a cofactor regenerating system). Comparisons of enzyme activities are made using a defined preparation of enzyme, a defined assay under a set condition, and one or more defined substrates, as further described in detail herein. Generally, when enzymes in cell lysates are compared, the numbers of cells and the amount of protein assayed are determined as well as use of identical expression systems and identical host cells to minimize variations in amount of enzyme produced by the host cells and present in the lysates.

[0051] "Conversion" refers to the enzymatic transformation of a substrate to the corresponding product. "Percent conversion" refers to the percent of the substrate that is converted to the product within a period of time under specified conditions. Thus, for example, the "enzymatic activity" or "activity" of a ketoreductase polypeptide can be expressed as "percent conversion" of the substrate to the product.

[0052] "Thermostable" or "thermal stable" are used interchangeably to refer to a polypeptide that is resistant to inactivation when exposed to a set of temperature conditions (e.g., 40-80°C) for a period of time (e.g., 0.5-24 hrs) compared to the untreated enzyme, thus retaining a certain level of residual activity (e.g., more than 60% to 80% for example) after exposure to elevated temperatures.

[0053] "Hydrophilic Amino Acid or Residue" refers to an amino acid or residue having a side chain exhibiting a hydrophobicity of less than zero according to the normalized consensus hydrophobicity scale of Eisenberg et al., 1984, J. Mol. Biol. 179:125-142. Genetically encoded hydrophilic amino acids include Thr (T), Ser (S), His (H), Glu (E), Asn (N), Gln (Q), Asp (D), Lys (K) and Arg (R).

[0054] "Acidic Amino Acid or Residue" refers to a hydrophilic amino acid or residue having a side chain exhibiting a pK value of less than about 6 when the amino acid is included in a peptide or polypeptide. Acidic amino acids typically have negatively charged side chains at physiological pH due to loss of a hydrogen ion. Genetically encoded acidic amino acids include Glu (E) and Asp (D).

[0055] "Basic Amino Acid or Residue" refers to a hydrophilic amino acid or residue having a side chain exhibiting a pK value of greater than about 6 when the amino acid is included in a peptide or polypeptide. Basic amino acids typically have positively charged side chains at physiological pH due to association with hydronium ion. Genetically encoded basic amino acids include Arg (R) and Lys (K).

[0056] "Polar Amino Acid or Residue" refers to a hydrophilic amino acid or residue having a side chain that is uncharged at physiological pH, but which has at least one bond in which the pair of electrons shared in common by two atoms is held more closely by one of the atoms. Genetically encoded polar amino acids include Asn (N), Gln (Q), Ser (S) and Thr (T).

[0057] "Hydrophobic Amino Acid or Residue" refers to an amino acid or residue having a side chain exhibiting a hydrophobicity of greater than zero according to the normalized consensus hydrophobicity scale of Eisenberg et al., 1984, J. Mol. Biol. 179:125-142. Genetically encoded hydrophobic amino acids include Pro (P), Ile (I), Phe (F), Val (V), Leu (L), Trp (W), Met (M), Ala (A) and Tyr (Y).

[0058] "Aromatic Amino Acid or Residue" refers to a hydrophilic or hydrophobic amino acid or residue having a side chain that includes at least one aromatic or heteroaromatic ring. Genetically encoded aromatic amino acids include Phe (F), Tyr (Y) and Trp (W). Although owing to its heteroaromatic ring side chain His (H) is classified as an aromatic residue, it may also be classified as a basic residue owing to pKa of its heteroaromatic nitrogen atom.

[0059] "Non-polar Amino Acid or Residue" refers to a hydrophobic amino acid or residue having a side chain that is uncharged at physiological pH and which has bonds in which the pair of electrons shared in common by two atoms is generally held equally by each of the two atoms (i.e., the side chain is not polar). Genetically encoded non-polar amino acids include Gly (G), Leu (L), Val (V), Ile (I), Met (M) and Ala (A).

[0060] "Aliphatic Amino Acid or Residue" refers to a hydrophobic amino acid or residue having an aliphatic hydrocarbon side chain. Genetically encoded aliphatic amino acids include Ala (A), Val (V), Leu (L) and Ile (I).

[0061] The amino acid Cys (C) is unique in that it can form disulfide bridges with other Cys (C) amino acids or other sulfanyl- or sulfhydryl-containing amino acids. The ability of Cys (and other amino acids with -SH containing side chains) to exist in a polypeptide in either the reduced free -SH or oxidized disulfide-bridged form affects whether it contributes net hydrophobic or hydrophilic character to the polypeptide. While Cys exhibits a hydrophobicity of 0.29 according to the normalized consensus scale of Eisenberg (Eisenberg et al., 1984, supra), it is to be understood that for purposes of the present disclosure, Cys is classified into its own unique group.

[0062] The amino acid Pro (P) has a conformationally constrained nature. Although it has hydrophobic properties, as used herein, Pro (P) or other similar residues is classified as a "conformationally constrained."

[0063] "Hydroxyl-containing Amino Acid or Residue" refers to an amino acid or residue containing a hydroxyl (-OH) moiety. Genetically-encoded hydroxyl-containing amino acids include Ser (S) and Thr (T). While L-Tyr (Y) contains a hydroxyl moiety, it is classified herein as an aromatic amino acid or residue.

[0064] "Amino acid difference" or "residue difference" refers to a change in the residue at a specified position of a polypeptide sequence when compared to a reference sequence. For example, a residue difference at position X3, where

the reference sequence has a glutamine, refers to a change of the residue at position X3 to any residue other than glutamine. As disclosed herein, an enzyme can include one or more residue differences relative to a reference sequence, where multiple residue differences typically are indicated by a list of the specified positions where changes are made relative to the reference sequence. The residue differences can be non-conservative changes or conservative changes.

In some embodiments, the residue differences can be conservative substitutions, non-conservative substitutions, or a combination of non-conservative and conservative substitutions. For the descriptions of the non-naturally occurring polypeptides herein, the amino acid residue position in the reference sequence is determined in the ketoreductase polypeptide beginning from the initiating methionine (M) residue (i.e., M represents residue position 1), although it will be understood by the skilled artisan that this initiating methionine residue may be removed by biological processing machinery, such as in a host cell or in vitro translation system, to generate a mature protein lacking the initiating methionine residue. The polypeptide sequence position at which a particular amino acid or amino acid change ("residue difference") is present is sometimes described herein as "Xn", or "position n", where n refers to the residue position with respect to the reference sequence. Where applicable, a specific substitution mutation, which is a replacement of the specific residue in a reference sequence with a different specified residue may be denoted by the conventional notation "X(number)Y", where X is the single letter identifier of the residue in the reference sequence, "number" is the residue position in the reference sequence, and Y is the single letter identifier of the residue substitution in the engineered sequence.

[0065] "Conservative amino acid substitutions" refer to the interchangeability of residues having similar side chains, and thus typically involves substitution of the amino acid in the polypeptide with amino acids within the same or similar defined class of amino acids. By way of example and not limitation, an amino acid with an aliphatic side chain may be substituted with another aliphatic amino acid, e.g., alanine, valine, leucine, and isoleucine; an amino acid with hydroxyl side chain is substituted with another amino acid with a hydroxyl side chain, e.g., serine and threonine; an amino acids having aromatic side chains is substituted with another amino acid having an aromatic side chain, e.g., phenylalanine, tyrosine, tryptophan, and histidine; an amino acid with a basic side chain is substituted with another amino acid with a basis side chain, e.g., lysine and arginine; an amino acid with an acidic side chain is substituted with another amino acid with an acidic side chain, e.g., aspartic acid or glutamic acid; and a hydrophobic or hydrophilic amino acid is replaced with another hydrophobic or hydrophilic amino acid, respectively. Exemplary conservative substitutions are provided in Table 1.

Table 1

Residue	Possible Conservative Substitutions
A, L, V, I	Other aliphatic (A, L, V, I) Other non-polar (A, L, V, I, G, M)
G, M	Other non-polar (A, L, V, I, G, M)
D, E	Other acidic (D, E)
K, R	Other basic (K, R)
N, Q, S, T	Other polar
H, Y, W, F	Other aromatic (H, Y, W, F)
C, P	None

[0066] "Non-conservative substitution" refers to substitution of an amino acid in the polypeptide with an amino acid with significantly differing side chain properties. Non-conservative substitutions may use amino acids between, rather than within, the defined groups and affects (a) the structure of the peptide backbone in the area of the substitution (e.g., proline for glycine) (b) the charge or hydrophobicity, or (c) the bulk of the side chain. By way of example and not limitation, an exemplary non-conservative substitution can be an acidic amino acid substituted with a basic or aliphatic amino acid; an aromatic amino acid substituted with a small amino acid; and a hydrophilic amino acid substituted with a hydrophobic amino acid.

[0067] "Deletion" refers to modification of the polypeptide by removal of one or more amino acids from the reference polypeptide. Deletions can comprise removal of 1 or more amino acids, 2 or more amino acids, 5 or more amino acids, 10 or more amino acids, 15 or more amino acids, or 20 or more amino acids, up to 10% of the total number of amino acids, or up to 20% of the total number of amino acids making up the polypeptide while retaining enzymatic activity and/or retaining the improved properties of an engineered ketoreductase enzyme. Deletions can be directed to the internal portions and/or terminal portions of the polypeptide. In various embodiments, the deletion can comprise a continuous segment or can be discontinuous.

[0068] "Insertion" refers to modification of the polypeptide by addition of one or more amino acids to the reference polypeptide. In some embodiments, the improved engineered ketoreductase enzymes comprise insertions of one or more amino acids to the naturally occurring ketoreductase polypeptide as well as insertions of one or more amino acids to other improved ketoreductase polypeptides. Insertions can be in the internal portions of the polypeptide, or to the carboxy or amino terminus. Insertions as used herein include fusion proteins as is known in the art. The insertion can be a contiguous segment of amino acids or separated by one or more of the amino acids in the naturally occurring polypeptide.

[0069] "Fragment" as used herein refers to a polypeptide that has an amino-terminal and/or carboxy-terminal deletion, but where the remaining amino acid sequence is identical to the corresponding positions in the sequence. Fragments can typically have about 80%, 90%, 95%, 98%, and 99% of the full-length ketoreductase polypeptide, for example the polypeptide of SEQ ID NO:2.

[0070] "Isolated polypeptide" refers to a polypeptide which is substantially separated from other contaminants that naturally accompany it, *e.g.*, protein, lipids, and polynucleotides. The term embraces polypeptides which have been removed or purified from their naturally-occurring environment or expression system (*e.g.*, host cell or *in vitro* synthesis). The improved ketoreductase enzymes may be present within a cell, present in the cellular medium, or prepared in various forms, such as lysates or isolated preparations. As such, in some embodiments, the engineered ketoreductase polypeptides of the present disclosure can be an isolated polypeptide.

[0071] "Substantially pure polypeptide" refers to a composition in which the polypeptide species is the predominant species present (*i.e.*, on a molar or weight basis it is more abundant than any other individual macromolecular species in the composition), and is generally a substantially purified composition when the object species comprises at least about 50 percent of the macromolecular species present by mole or % weight. Generally, a substantially pure engineered ketoreductase polypeptide composition will comprise about 60 % or more, about 70% or more, about 80% or more, about 90% or more, about 95% or more, and about 98% or more of all macromolecular species by mole or % weight present in the composition. Solvent species, small molecules (<500 Daltons), and elemental ion species are not considered macromolecular species. In some embodiments, the isolated improved ketoreductase polypeptide is a substantially pure polypeptide composition.

[0072] "Heterologous" polynucleotide refers to any polynucleotide that is introduced into a host cell by laboratory techniques, and includes polynucleotides that are removed from a host cell, subjected to laboratory manipulation, and then reintroduced into a host cell.

[0073] "Codon optimized" refers to changes in the codons of the polynucleotide encoding a protein to those preferentially used in a particular organism such that the encoded protein is efficiently expressed in the organism of interest. In some embodiments, the polynucleotides encoding the ketoreductase enzymes may be codon optimized for optimal production from the host organism selected for expression.

[0074] "Control sequence" is defined herein to include all components, which are necessary or advantageous for the expression of a polynucleotide and/or polypeptide of the present disclosure. Each control sequence may be native or foreign to the polynucleotide of interest. Such control sequences include, but are not limited to, a leader, polyadenylation sequence, propeptide sequence, promoter, signal peptide sequence, and transcription terminator.

[0075] "Operably linked" is defined herein as a configuration in which a control sequence is appropriately placed (*i.e.*, in a functional relationship) at a position relative to a polynucleotide of interest such that the control sequence directs or regulates the expression of the polynucleotide and/or polypeptide of interest.

[0076] "Cofactor regeneration system" refers to a set of reactants that participate in a reaction that reduces the oxidized form of the cofactor (*e.g.*, NADP⁺ to NADPH). Cofactors oxidized by the ketoreductase-catalyzed reduction of the keto substrate are regenerated in reduced form by the cofactor regeneration system. Cofactor regeneration systems comprise a stoichiometric reductant that is a source of reducing hydrogen equivalents and is capable of reducing the oxidized form of the cofactor. The cofactor regeneration system may further comprise a catalyst, for example an enzyme catalyst that catalyzes the reduction of the oxidized form of the cofactor by the reductant. Cofactor regeneration systems to regenerate NADH or NADPH from NAD⁺ or NADP⁺, respectively, are known in the art and may be used in the methods described herein.

[0077] The terms "glucose dehydrogenase" and "GDH" are used interchangeably herein to refer to an NAD⁺ or NADP⁺-dependent enzyme that catalyzes the conversion of D-glucose and NAD⁺ or NADP⁺ to gluconic acid and NADH or NADPH, respectively.

[0078] The term "secondary alcohol dehydrogenase" is used herein to refer to an NAD⁺ or NADP⁺-dependent enzyme that catalyzes the conversion of a secondary alcohol (*e.g.*, isopropanol) and NAD⁺ or NADP⁺ to a ketone and NADH or NADPH, respectively.

[0079] "Alkyl" refers to groups of from 1 to 12 carbon atoms inclusively, either straight chained or branched, more preferably from 1 to 8 carbon atoms inclusively, and most preferably 1 to 6 carbon atoms inclusively.

[0080] "Alkenyl" refers to groups of from 2 to 12 carbon atoms inclusively, either straight or branched containing at least one double bond but optionally containing more than one double bond.

[0081] "Alkynyl" refers to groups of from 2 to 12 carbon atoms inclusively, either straight or branched containing at least one triple bond but optionally containing more than one triple bond, and additionally optionally containing one or more double bonded moieties.

[0082] "Alkoxy" refers to the group alkyl-O- wherein the alkyl group is as defined above including optionally substituted alkyl groups as also defined above.

[0083] "Alkenoxy" refers to the group alkenyl-O- wherein the alkenyl group is as defined above including optionally substituted alkenyl groups as also defined above.

[0084] "Alkynoxy" refers to the group alkynyl-O- wherein the alkynyl group is as defined above including optionally substituted alkynyl groups as also defined above.

[0085] "Aryl" refers to an unsaturated aromatic carbocyclic group of from 6 to 14 carbon atoms inclusively having a single ring (e.g., phenyl) or multiple condensed rings (e.g., naphthyl or anthryl). Preferred aryls include phenyl, naphthyl and the like.

[0086] "Arylalkyl" refers to aryl -alkyl- groups preferably having from 1 to 6 carbon atoms inclusively in the alkyl moiety and from 6 to 10 carbon atoms inclusively in the aryl moiety. Such arylalkyl groups are exemplified by benzyl, phenethyl and the like.

[0087] "Arylalkenyl" refers to aryl -alkenyl- groups preferably having from 2 to 6 carbon atoms in the alkenyl moiety and from 6 to 10 carbon atoms inclusively in the aryl moiety.

[0088] "Arylalkynyl" refers to aryl -alkynyl- groups preferably having from 2 to 6 carbon atoms inclusively in the alkynyl moiety and from 6 to 10 carbon atoms inclusively in the aryl moiety.

[0089] "Cycloalkyl" refers to cyclic alkyl groups of from 3 to 12 carbon atoms inclusively having a single cyclic ring or multiple condensed rings which can be optionally substituted with from 1 to 3 alkyl groups. such cycloalkyl groups include, by way of example, single ring structures such as cyclopropyl, cyclobutyl, cyclopentyl, cyclooctyl, 1-methylcyclopropyl, 2-methylcyclopentyl, 2-methylcyclooctyl, and the like, or multiple ring structures such as adamantyl, and the like.

[0090] "Cycloalkenyl" refers to cyclic alkenyl groups of from 4 to 12 carbon atoms inclusively having a single cyclic ring or multiple condensed rings and at least one point of internal unsaturation, which can be optionally substituted with from 1 to 3 alkyl groups. examples of suitable cycloalkenyl groups include, for instance, cyclobut-2-enyl, cyclopent-3-enyl, cyclooct-3-enyl and the like.

[0091] "Cycloalkylalkyl" refers to cycloalkyl -alkyl- groups preferably having from 1 to 6 carbon atoms inclusively in the alkyl moiety and from 6 to 10 carbon atoms inclusively in the cycloalkyl moiety. Such cycloalkylalkyl groups are exemplified by cyclopropylmethyl, cyclohexylethyl and the like.

[0092] "Cycloalkylalkenyl" refers to cycloalkyl -alkenyl- groups preferably having from 2 to 6 carbon atoms inclusively in the alkenyl moiety and from 6 to 10 carbon atoms inclusively in the cycloalkyl moiety. Such cycloalkylalkenyl groups are exemplified by cyclohexylethenyl and the like.

[0093] "Cycloalkylalkynyl" refers to cycloalkyl -alkynyl- groups preferably having from 2 to 6 carbon atoms inclusively in the alkynyl moiety and from 6 to 10 carbon atoms inclusively in the cycloalkyl moiety. Such cycloalkylalkynyl groups are exemplified by cyclopropylethynyl and the like.

[0094] "Heteroaryl" refers to a monovalent aromatic heterocyclic group of from 1 to 10 carbon atoms inclusively and 1 to 4 heteroatoms inclusively selected from oxygen, nitrogen and sulfur within the ring. Such heteroaryl groups can have a single ring (e.g., pyridyl or furyl) or multiple condensed rings (e.g., indolizynyl or benzothienyl).

[0095] "Heteroarylalkyl" refers to heteroaryl -alkyl- groups preferably having from 1 to 6 carbon atoms inclusively in the alkyl moiety and from 6 to 10 atoms inclusively in the heteroaryl moiety. Such heteroarylalkyl groups are exemplified by pyridylmethyl and the like.

[0096] "Heteroarylalkenyl" refers to heteroaryl -alkenyl- groups preferably having from 2 to 6 carbon atoms inclusively in the alkenyl moiety and from 6 to 10 atoms inclusively in the heteroaryl moiety.

[0097] "Heteroarylalkynyl" refers to heteroaryl -alkynyl- groups preferably having from 2 to 6 carbon atoms inclusively in the alkynyl moiety and from 6 to 10 atoms inclusively in the heteroaryl moiety.

[0098] "Heterocycle" refers to a saturated or unsaturated group having a single ring or multiple condensed rings, from 1 to 8 carbon atoms inclusively and from 1 to 4 hetero atoms inclusively selected from nitrogen, sulfur or oxygen within the ring. such heterocyclic groups can have a single ring (e.g., piperidiny or tetrahydrofuryl) or multiple condensed rings (e.g., indoliny, dihydrobenzofuran or quinuclidiny). preferred heterocycles include piperidiny, pyrrolidiny and tetrahydrofuryl.

[0099] Examples of heterocycles and heteroaryls include, but are not limited to, furan, thiophene, thiazole, oxazole, pyrrole, imidazole, pyrazole, pyridine, pyrazine, pyrimidine, pyridazine, indolizine, isoindole, indole, indazole, purine, quinolizine, isoquinoline, quinoline, phthalazine, naphthylpyridine, quinoxaline, quinazoline, cinnoline, pteridine, carbazole, carboline, phenanthridine, acridine, phenanthroline, isothiazole, phenazine, isoxazole, phenoxazine, phenothiazine, imidazolidine, imidazoline, piperidine, piperazine, pyrrolidine, indoline and the like.

[0100] Unless otherwise specified, positions occupied by hydrogen in the foregoing groups can be further substituted with substituents exemplified by, but not limited to, hydroxy, oxo, nitro, methoxy, ethoxy, alkoxy, substituted alkoxy,

trifluoromethoxy, haloalkoxy, fluoro, chloro, bromo, iodo, halo, methyl, ethyl, propyl, butyl, alkyl, alkenyl, alkynyl, substituted alkyl, trifluoromethyl, haloalkyl, hydroxyalkyl, alkoxyalkyl, thio, alkylthio, acyl, carboxy, alkoxy carbonyl, carboxamido, substituted carboxamido, alkylsulfonyl, alkylsulfinyl, alkylsulfonylamino, sulfonamido, substituted sulfonamido, cyano, amino, substituted amino, alkylamino, dialkylamino, aminoalkyl, acylamino, amidino, amidoximo, hydroxamoyl, phenyl, aryl, substituted aryl, aryloxy, arylalkyl, arylalkenyl, arylalkynyl, pyridyl, imidazolyl, heteroaryl, substituted heteroaryl, heteroaryloxy, heteroarylalkyl, heteroarylalkenyl, heteroarylalkynyl, cyclopropyl, cyclobutyl, cyclopentyl, cyclohexyl, cycloalkyl, cycloalkenyl, cycloalkylalkyl, substituted cycloalkyl, cycloalkyloxy, pyrrolidinyl, piperidinyl, morpholino, heterocycle, (heterocycle)oxy, and (heterocycle)alkyl; and preferred heteroatoms are oxygen, nitrogen, and sulfur. It is understood that where open valences exist on these substituents they can be further substituted with alkyl, cycloalkyl, aryl, heteroaryl, and/or heterocycle groups, that where these open valences exist on carbon they can be further substituted by halogen and by oxygen-, nitrogen-, or sulfur-bonded substituents, and where multiple such open valences exist, these groups can be joined to form a ring, either by direct formation of a bond or by formation of bonds to a new heteroatom, preferably oxygen, nitrogen, or sulfur. It is further understood that the above substitutions can be made provided that replacing the hydrogen with the substituent does not introduce unacceptable instability to the molecules of the present invention, and is otherwise chemically reasonable.

[0101] The term "suitable reaction conditions" refers to those conditions in the biocatalytic reaction solution (e.g., ranges of enzyme loading, substrate loading, cofactor loading, T, pH, buffers, cosolvents, etc.) under which a non-naturally occurring ketoreductase polypeptide of the present disclosure is capable of converting compound (2) to compound (1) (or compound of Formula IIa to compound of Formula Ia). Exemplary "suitable reaction conditions" are provided in the present disclosure and illustrated by the Examples.

4.3 Non-Naturally Occurring or Engineered Ketoreductase Polypeptides

[0102] Enzymes belonging to the ketoreductase (KRED) or carbonyl reductase class (EC1.1.1.184) are useful for the synthesis of optically active alcohols from the corresponding prostereoisomeric ketone substrates and by stereospecific reduction of corresponding racemic aldehyde and ketone substrates. KREDs typically convert a ketone or aldehyde substrate to the corresponding alcohol product, but may also catalyze the reverse reaction, oxidation of an alcohol substrate to the corresponding ketone/aldehyde product. The reduction of ketones and aldehydes and the oxidation of alcohols by enzymes such as KRED requires a co-factor, most commonly reduced nicotinamide adenine dinucleotide (NADH) or reduced nicotinamide adenine dinucleotide phosphate (NADPH), and nicotinamide adenine dinucleotide (NAD) or nicotinamide adenine dinucleotide phosphate (NADP+) for the oxidation reaction. NADH and NADPH serve as electron donors, while NAD and NADP serve as electron acceptors. It is frequently observed that KREDs and other alcohol dehydrogenases accept either the phosphorylated or the non-phosphorylated co-factor (in its oxidized and reduced state).

[0103] KREDs are being used increasingly in place of chemical procedures for the conversion of different keto and aldehyde compounds to chiral alcohol products. These biocatalytic conversions can employ whole cells expressing the ketoreductase for biocatalytic ketone reductions, or purified enzymes, particularly in those instances where presence of multiple ketoreductases in whole cells would adversely affect the enantiomeric purity and yield of the desired product. For *in vitro* applications, a co-factor (NADH or NADPH) regenerating enzyme such as glucose dehydrogenase (GDH), formate dehydrogenase *etc.* is used in conjunction with the ketoreductase.

[0104] Examples illustrating the use of naturally occurring or engineered KREDs in biocatalytic processes to generate useful chemical compounds include asymmetric reduction of 4-chloroacetoacetate esters (Zhou, J. Am. Chem. Soc. 1983 105:5925-5926; Santaniello, J. Chem. Res. (S) 1984:132-133; U.S. Patent Nos. 5,559,030; U.S. Patent No. 5,700,670 and U.S. Patent No. 5,891,685), reduction of dioxocarboxylic acids (e.g., U.S. Patent No. 6,399,339), reduction of tert-butyl (S)-chloro-5-hydroxy-3-oxohexanoate (e.g., U.S. Patent No. 6,645,746 and WO 01/40450), reduction pyrrolotriazine-based compounds (e.g., U.S. application No. 2006/0286646); reduction of substituted acetophenones (e.g., U.S. Patent No. 6,800,477); and reduction of ketothiolanes (WO 2005/054491).

[0105] KREDs can be found in a wide range of bacteria and yeasts (for reviews: Kraus and Waldman, Enzyme catalysis in organic synthesis Vols. 1&2. VCH Weinheim 1995; Faber, K., Biotransformations in organic chemistry, 4th Ed. Springer, Berlin Heidelberg New York. 2000; Hummel and Kula Eur. J. Biochem. 1989 184:1-13). Several KRED gene and enzyme sequences have been reported, including: *Candida magnoliae* (Genbank Acc. No. JC7338; GI:11360538); *Candida parapsilosis* (Genbank Acc. No. BAA24528.1; GI:2815409), *Sporobolomyces salmonicolor* (Genbank Acc. No. AF160799; GI:6539734); *Lactobacillus kefir* (Genbank Acc. No. AAP94029.1; GI: 33112056); *Lactobacillus brevis* (Genbank Acc. No. 1NXQ_A; GI: 30749782); and *Thermoanaerobium brockii* (Genbank Acc. No. P 14941; GI: 1771790).

[0106] These naturally occurring ketoreductase polypeptides do not efficiently convert compound (2) to compound (1). The non-naturally occurring, engineered, ketoreductase polypeptides of the present disclosure, however, are capable of carrying out this conversion with improved properties including, high diastereomeric excess (e.g., at least about 99% d.e.), increased activity (e.g., at least about 10-fold increased activity relative to the reference polypeptide SEQ ID NO:2),

high percent conversion (e.g., at least about 90% conversion in 24 h), in the presence of high substrate loadings (e.g., at least about 50 g/L compound (2)), and without any cofactor regenerating enzyme other than the engineered ketoreductase polypeptide.

[0107] The non-naturally occurring polypeptides of the present disclosure are synthetic variants of the naturally occurring ketoreductase of *Lactobacillus kefir*, and comprise amino acid sequences that have one or more residue differences as compared to the reference sequence of the synthetic variant ketoreductase polypeptide of SEQ ID NO:2. The residue differences occur at residue positions that affect enzyme activity, stereoselectivity, thermostability, expression, or various combinations thereof. The residue differences provide structural changes that allow the engineered polypeptides to convert the ketophenol substrate, 1-(4-fluorophenyl)-3(R)-[3-(4-fluorophenyl)-3-oxopropyl]-4(S)-(4-hydroxyphenyl)-2-azetidinone (compound (2); MW 407.41) to the chiral alcohol product, 1-(4-fluorophenyl)-3(R)-[3-(4-fluorophenyl)-3(S)-hydroxypropyl]-4(S)-(4-hydroxyphenyl)-2-azetidinone (compound (1); MW 409.43) (as illustrated in Scheme 1) with at least 2-fold, at least 10-fold, at least 25-fold, at least 40-fold, or at least 60-fold increased activity relative to the activity of the reference polypeptide of SEQ ID NO: 2. Further these engineered polypeptides are capable of highly stereoselective conversion of compound (2) to compound (1) in at least about 97%, about 98%, or at least about 99% diastereomeric excess. Further, in some embodiments these non-naturally occurring polypeptides are capable of catalyzing the conversion of compound (2) to compound (1) using added cofactor (NADPH or NADH), or in presence of a co-factor recycling system, for example an appropriate dehydrogenase (e.g., glucose dehydrogenase, formate dehydrogenase or ketoreductase/alcohol dehydrogenase) and a suitable dehydrogenase substrate, such as glucose, glucose-6-phosphate, formate, or a secondary alcohol, e.g., isopropanol. In some embodiments, the non-naturally occurring ketoreductase polypeptides can function not only to convert compound (2) to compound (1), but also function as the secondary alcohol dehydrogenase of a cofactor recycling system and thereby recycle the cofactor in the presence of a secondary alcohol. Thus, the engineered biocatalysts present disclosure are capable of providing highly efficient biocatalytic processes for preparing Ezetimibe as substantially enantiomerically pure preparations.

[0108] Structure and function information for exemplary non-naturally occurring (or engineered) ketoreductase polypeptides of the present disclosure are shown below in Table 2. The odd numbered sequence identifiers (i.e., SEQ ID NOs) refer to the nucleotide sequence encoding the amino acid sequence provided by the even numbered SEQ ID NOs, and the sequences are provided in the electronic sequence listing file accompanying this disclosure. The amino acid residue differences are based on comparison to the reference sequence of SEQ ID NO: 2, which is an engineered ketoreductase having the following 19 residue differences relative to the amino acid sequence of the naturally occurring wild-type ketoreductase of *Lactobacillus kefir* (Genbank acc. No. AAP94029.1; GI: 33112056): D3N, G7S, L17Q, V95L, S96Q, G117S, Q127R, E145S, F147L, T152M, L153V, L176V, Y190C, D198K, L199D, E200P, K211R, I223V, and A241S. The activity of each engineered polypeptide relative to the reference polypeptide of SEQ ID NO: 2 was determined as conversion of substrate of compound (2) to product of compound (1) over a 24 h period at room temperature in a 96-well plate format assay of cell lysates containing the engineered polypeptides. General assay protocol and reaction conditions were as follows (with exceptions noted in Table 2): 60 μ L of a 13.33 g/L solution of the compound (2) in toluene:IPA:acetone (v/v/v ratio of 5:9:1) added to each well of a Costar™ deep-well 96-well plate; subsequently, 120 μ L of a 0.8 g/L solution of NADP in 100 mM TEA buffer, pH 7.0 containing 1 mM MgSO₄ added to each well; finally, 20 μ L of a freshly prepared suspension of lysed cells (i.e., cells expressing the variant polypeptide) in lysis buffer were added to make the total volume in each well 200 μ L. Final conditions in each well (except as noted below): [compound (2)] = 4 g/L, [NADP] = 0.5 g/L, Solvent = toluene:IPA:acetone:buffer (relative % volumes of 10:18:2:70). The plate was then heat sealed and shaken for 24 h at RT (or 37 °C) before 0.8 mL of acetonitrile was added to each well to quench the reaction. The levels of activity (i.e., "+" "++" "+++" etc.) are defined as follows: "+" indicates at least equal to but less than 2 times the activity of SEQ ID NO: 2; "++" indicates at least 2 times but less than 10 times the activity of SEQ ID NO: 2; "+++" indicates at least 10 times but less than 25 times the activity of SEQ ID NO: 2; "++++" indicates at least 25 times but less than 40 times the activity of SEQ ID NO: 2; "+++++" indicates at least 40 times but less than 60 times the activity of SEQ ID NO: 2; "++++++" indicates at least 60 times the activity of SEQ ID NO: 2.

Table 2

SEQ ID NO	Residue differences	Activity
(nt/aa)	(relative to SEQ ID NO:2)	(relative to SEQ ID NO: 2)
3/4	H40R; V153I; C190A; E204V;	++
5/6	H40R; I93A; A94T; V153I; C190A; V196T; D199F; M206I;	++
7/8	H40R;	+
9/10	H40R; V148I;	+

EP 2 566 497 B1

(continued)

	SEQ ID NO	Residue differences	Activity
	(nt/aa)	(relative to SEQ ID NO:2)	(relative to SEQ ID NO: 2)
5	11/12	H40R; V196A;	+
	13/14	H40R; S207T;	++
	15/16	H40R; Q96V;	+
10	17/18	H40R; V196S;	+
	19/20	H40R; V196C;	++
	21/22	H40R; V196N;	++
15	23/24	H40R; A202L;	++
	25/26	H40R; A202N;	+
	27/28	H40R; A202V;	++
20	29/30	H40R; I93A; A94P; V153I; C190A; V196T; D199F; M206I;	++
	31/32	H40R; Q96G; V153I; C190A; V196T; D199F; M206I;	++
25	33/34	H40R; Q96A; V153I; C190A; V196T; D199F; M206I;	++
	35/36	H40R; I93A; A94T; V153I; C190A; V196T; D199F; E203G; M206I;	++
30	37/38	H40R; I93A; A94T; Q96V; V153I; C190A; V196T; D199F; M206I;	+
	39/40	H40R; I93A; A94T; V153I; C190A; L195M; V196T; D199F; M206I;	++
35	41/42	H40R; I93A; A94T; V153I; C190A; V196T; D199F; M206I; S207C;	++
	43/44	H40R; I93A; A94T; V153I; C190A; V196T; D199F; M206I; S207I;	++
40	45/46	H40R; I93A; A94T; V153I; C190A; V196T; D199F; M206I; S207N;	++
	47/48	L21R; D25R; H40R; I93A; A94T; S117G; R127K; L147M; V153I; C190A; V196T; D199F; E203G; M206I; N221D;	+++ ¹
45	49/50	L21F; D25R; H40R; I93A; A94T; S117G; L147M; V153I; C190A; V196T; D199F; E203G; M206I; N221D;	+++ ¹
	51/52	H40R; I93A; A94S; Q96N; V153L; C190A; V196T; D199F; A202G; E203G; M206I;	+++ ¹
50	53/54	D25T; H40R; I93A; A94T; R108H; V153I; C190A; V196T; D199F; E203G; M206I;	+++ ¹
	55/56	H40R; I93A; A94T; D150H; V153I; C190A; V196T; D199F; E203G; M206I;	+++ ¹
55	57/58	H40R; I93A; A94T; V153I; C190P; V196T; D199F; E203G; M206I;	+++ ¹

EP 2 566 497 B1

(continued)

	SEQ ID NO	Residue differences	Activity
	(nt/aa)	(relative to SEQ ID NO:2)	(relative to SEQ ID NO: 2)
5	59/60	H40R; I93A; A94T; V153I; C190A; V196T; D199F; E203G; E204A; M206I;	+++ ¹
	61/62	H40R; I93A; A94T; V153I; C190A; V196T; D199F; E203G; E204V; M206I;	+++ ¹
10	63/64	H40R; I93A; A94T; V153I; C190A; V196T; D199F; E203G; M205V; M206I;	+++ ¹
	65/66	H40R; I93A; A94T; M152N; V153I; C190A; V196T; D199F; E203G; M206I;	+++ ¹
15	67/68	H40R; I93A; A94T; M152F; V153I; C190A; V196T; D199F; E203G; M206I;	+++ ¹
	69/70	H40R; I93A; A94T; Q96A; V153I; C190A; V196T; D199F; E203G; M206I; S207T;	+++ ¹
20	71/72	D25N; H40R; I93A; A94T; V153I; C190A; V196T; D199F; E203G; M206I;	+++ ¹
	73/74	D25T; H40R; I93A; A94T; V153I; C190A; V196T; D199F; E203G; M206I;	+++ ¹
25	75/76	H40R; I93A; A94T; V153I; C190A; L195M; V196T; D199F; E203G; M206I; S207C;	+++ ¹
	77/78	H40R; I93A; A94S; Q96S; V153L; C190A; V196T; D199F; A202G; E203G; M206I;	+++ ¹
30	79/80	H40R; I93A; A94S; Q96P; V153L; C190A; V196T; D199F; A202G; E203G; M206I;	+++ ¹
	81/82	H40R; I93A; A94S; Q96A; V153L; C190A; V196T; D199F; A202G; E203G; M206I;	++++ ¹
35	83/84	H40R; I93A; A94S; Q96T; V153L; C190A; V196T; D199F; A202G; E203G; M206I;	+++ ¹
	85/86	H40R; I93A; A94S; Q96N; V153L; C190A; V196T; D199F; A202G; E203G; M206I; R211K;	+++ ¹
40	87/88	H40R; I93A; A94S; Q96N; V153L; A155C; C190A; V196T; D199F; A202G; E203G; M206I;	+++ ¹
	89/90	H40R; I93A; A94S; Q96N; V153L; C190A; V196T; D199F; G201I; A202G; E203G; M206I;	+++ ¹
45	91/92	H40R; I93A; A94S; Q96N; V153L; C190A; V196T; D199F; G201L; A202G; E203G; M206I;	+++ ¹
	93/94	H40R; I93A; A94S; Q96N; V153L; C190A; V196T; D199F; G201A; A202G; E203G; M206I;	+++ ¹
50	95/96	H40R; I93A; A94S; L95V; Q96N; R127Q; V153L; C190A; V196T; D199F; A202G; E203G; M206I; R211K;	+++ ¹
	97/98	H40R; A94S; L95M; Q96N; L147M; V153L; C190A; V196T; D199F; G201I; A202G; E203G; M206I;	+++ ¹
55	99/100	H40R; A94S; Q96P; V153L; C190A; V196T; D199F; G201A; A202G; E203G; M206I; R211K;	++++ ²

EP 2 566 497 B1

(continued)

	SEQ ID NO	Residue differences	Activity
	(nt/aa)	(relative to SEQ ID NO:2)	(relative to SEQ ID NO: 2)
5	101/102	L21F; D25R; H40R; I93A; A94S; Q96P; R108K; S117G; L147M; V153L; C190A; V196T; D199F; A202G; E203G; M206I; N221D;	++++ ²
10	103/104	L21R; D25R; H40R; I93A; A94S; Q96P; R108K; S117G; L147M; V153L; C190A; V196T; D199F; A202G; E203G; M206I; N221D;	++++ ²
15	105/106	L21F; D25R; H40R; I93A; A94S; Q96P; R108K; R127K; L147M; V153L; C190A; V196T; D199F; A202G; E203G; M206I; N221D;	++++ ²
20	107/108	L21R; D25R; H40R; I93A; A94S; Q96P; R108K; S117G; L147M; V153L; C190A; V196T; D199F; A202G; E203G; M206I;	++++ ²
25	109/110	L21F; D25R; H40R; I93A; A94S; Q96P; R108K; S117G; L147M; V153L; V163I; C190A; V196T; D199F; A202G; E203G; M206I; N221D;	++++ ²
30	111/112	L21R; H40R; I93A; A94S; Q96P; S117G; L147M; V153L; C190A; V196T; D199F; A202G; E203G; M206I; N221D;	++++ ²
35	113/114	L21R; H40R; I93A; A94S; Q96P; S117G; R127K; L147M; V153L; C190A; V196T; D199F; A202G; E203G; M206I; N221D;	++++ ²
40	115/116	H40R; I93A; A94S; Q96P; S117G; L147M; V153L; C190A; V196T; D199F; A202G; E203G; M206I; N221D;	++++ ²
45	117/118	D25R; H40R; I93A; A94S; Q96P; R108D; S117G; R127K; L147M; V153L; C190A; V196T; D199F; A202G; E203G; M206I; N221D;	++++ ²
50	119/120	H40R; I93A; A94S; Q96P; S117G; L147M; V153L; C190A; V196T; D199F; A202G; E203G; M206I;	++++ ²
55	121/122	H40R; I93A; A94S; Q96P; R108K; S117G; R127K; L147M; V153L; C190A; V196T; D199F; A202G; E203 G; M206I; N221D;	++++ ²
60	123/124	L21R; D25R; H40R; I93A; A94S; Q96P; S117G; R127K; V153L; C190A; V196T; D199F; A202G; E203G; M206I; N221D;	++++ ²
65	125/126	D25R; H40R; I93A; A94S; Q96P; S117G; R127K; V153L; C190A; V196T; D199F; A202G; E203G; M206I;	++++ ²
70	127/128	D25R; H40R; I93A; A94S; Q96P; R108K; S117G; R127K; L147M; V153L; C190A; V196T; D199F; A202G; E203G; M206I; N221D;	++++ ²
75	129/130	D25R; H40R; I93A; A94S; Q96P; L147M; V153L; C190A; V196T; D199F; A202G; E203G; M206I; N221D;	++++ ²

EP 2 566 497 B1

(continued)

	SEQ ID NO	Residue differences	Activity
	(nt/aa)	(relative to SEQ ID NO:2)	(relative to SEQ ID NO: 2)
5	131/132	L21R; D25R; H40R; I93A; A94S; Q96P; S117G; L147M; V153L; C190A; V196T; D199F; A202G; E203G; M206I; N221D;	++++ ²
10	133/134	H40R; I93T; A94S; Q96P; V153L; C190A; V196T; D199F; A202G; E203G; M206I;	++++ ²
	135/136	H40R; I93A; Q96P; V153L; C190A; V196T; D199F; A202G; E203G; M206I;	++++ ²
15	137/138	H40R; I93A; A94S; Q96P; V153L; C190A; V196T; D199F; A202G; E203G; M205V; M206I;	++++ ²
	139/140	H40R; A94S; Q96P; S117A; V153L; C190A; V196T; D199F; G201A; A202G; E203G; M206I; R211K;	++++ ²
20	141/142	H40R; A94S; Q96P; S117G; V153L; C190A; V196T; D199F; G201A; A202G; E203G; M206I; R211K;	++++ ³
	143/144	H40R; A94S; Q96P; V153L; C190A; V196T; D199W; G201A; A202G; E203G; M206I; R211K;	++++ ³
25	145/146	H40R; A94S; Q96P; L147M; V153L; C190A; V196T; D199F; G201A; A202G; E203G; M206I; R211K;	++++ ³
	147/148	H40R; A64V; A94S; Q96P; V153L; C190A; V196T; D199F; G201A; A202G; E203G; M206I; R211K;	++++ ³
30	149/150	H40R; A94S; Q96P; V153L; C190A; V196T; D199F; G201A; A202G; E203G; M206I; R211K; V223I;	++++ ³
35	151/152	H40R; A94S; Q96P; V153L; C190A; L195M; V196T; D199F; G201A; A202G; E203G; M206I; R211K;	++++ ³
	153/154	H40R; A94S; Q96P; L147I; V153L; C190A; V196T; D199F; G201A; A202G; E203G; M206I; R211K;	++++ ³
40	155/156	H40R; A94S; Q96P; V99L; R108H; L147I; V153L; C190A; V196T; D199F; G201A; A202G; E203G; M205V; M206I; R211K; I226V;	++++ ³
45	157/158	H40R; A94S; Q96P; V99L; R108H; L147I; V153L; C190A; V196T; D199F; G201A; A202G; E203G; M206I; R211K; I226V;	++++ ³
	159/160	H40R; A94S; Q96P; V99L; L147I; V153L; C190A; V196T; D199F; G201A; A202G; E203G; M205V; M206I; R211K;	++++ ³
50	161/162	H40R; A94S; Q96P; V99L; R108H; S117G; L147I; V153L; C190A; V196T; D199F; G201A; A202G; E203G; M206I; R211K; I226V;	++++ ³
55	163/164	H40R; A94S; Q96P; V99L; R108H; L147I; V153L; C190A; V196T; D199F; G201A; A202G; E203G; M205V; M206I; R211K;	++++ ³

(continued)

SEQ ID NO	Residue differences	Activity
(nt/aa)	(relative to SEQ ID NO:2)	(relative to SEQ ID NO: 2)
165/166	H40R; A94S; Q96P; V99L; L147I; V153L; C190A; V196T; D199F; G201A; A202G; E203G; M205V; M206I; R211K; I226V;	+++ ³
167/168	H40R; A64V; A94S; Q96P; V99L; R108H; S117G; L147I; V153L; C190A; V196T; D199W; G201A; A202G; E203G; M206I; R211K; I226V;	+++++ ⁴
¹ Modified final assay conditions: [compound (2)] = 4 g/L, [NADP] = 0.5 g/L, Solvent = toluene:IPA:acetone: buffer (100 mM TEA buffer, pH 7.0 containing 1 mM MgSO ₄) in relative % volumes of 30:18:2:50; plate shaken for 2 h at RT. ² Modified final assay conditions: [compound (2)] = 50 g/L, [NADP] = 0.1 g/L, Solvent = toluene:IPA:acetone:buffer (100 mM TEA buffer, pH 7.0 containing 1 mM MgSO ₄) in relative % volumes of 30:18:2:50; plate shaken for 24 h at 30°C. ³ Modified final assay conditions: [compound (2)] = 4 g/L, [NADP] = 0.1 g/L, Solvent = (IPA + acetone):buffer (100 mM TEA buffer, pH 7.0 containing 1 mM MgSO ₄) in relative % volumes of 70:30; plate shaken for 24 h at RT. ⁴ Modified final assay conditions: [compound (2)] = 80 g/L, [NADP] = 0.1 g/L, Solvent = (IPA + acetone):buffer (100 mM TEA buffer, pH 7.0 containing 1 mM MgSO ₄) in relative % volumes of 65:35; plate shaken for 24 h at RT.		

[0109] In some embodiments, the non-naturally occurring polypeptides of the present disclosure which are capable of converting compound (2) to compound (1) comprising an amino acid sequence selected from any one of SEQ ID NO: 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, or 168. As shown above in Table 2, each of these polypeptides comprises one or more amino acid residue differences as compared to SEQ ID NO: 2, and has at least 2-fold, at least 10-fold, at least 25-fold, at least 40-fold, or at least 60-fold increased activity relative to the activity of the reference polypeptide of SEQ ID NO: 2. Specific amino acid differences are shown in Table 2 and include one or more residue differences as compared to SEQ ID NO:2 at the following residue positions: **X21, X25, X40, X64, X93, X94, X95, X96, X99, X108, X117, X127, X147, X148, X150, X152, X153, X155, X190, X195, X196, X201, X202, X203, X204, X205, X206, X207, X211, X221, X223, and X226.**

[0110] In some embodiments, the present disclosure provides a non-naturally occurring polypeptide capable of converting compound (2) to compound (1) with at least 2-fold, at least 10-fold, at least 25-fold, at least 40-fold, or at least 60-fold increased activity relative to the activity of the polypeptide of SEQ ID NO: 2, and comprises an amino acid sequence having at least 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identity to a reference amino acid sequence selected from any one of SEQ ID NO: 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166; or 168.

[0111] In some embodiments, the present disclosure provides a non-naturally occurring polypeptide capable of converting compound (2) to compound (1) with at least 2-fold, at least 10-fold, at least 25-fold, at least 40-fold, or at least 60-fold increased activity relative to the activity of the polypeptide of SEQ ID NO: 2, comprises an amino acid sequence having at least 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identity to SEQ ID NO: 2, and further comprises a set of amino acid residue differences as compared to SEQ ID NO:2 of any one of SEQ ID NO: 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, or 168. In some embodiments, in addition to the set of amino acid residue differences of any one of the non-naturally occurring polypeptides of SEQ ID NO: 4 through SEQ ID NO: 168, the sequence of the non-naturally occurring polypeptide can further comprise 1-2, 1-3, 1-4, 1-5, 1-6, 1-7, 1-8, 1-9, 1-10, 1-11, 1-12, 1-14, 1-15, 1-16, 1-18, 1-20, 1-22, 1-24, 1-26, 1-30, 1-35, 1-40 residue differences at other amino acid residue positions as compared to the SEQ ID NO: 2. In some embodiments, the residue differences can comprise conservative substitutions and non-conservative substitutions as compared to SEQ ID NO: 2.

[0112] The present disclosure also contemplates a non-naturally occurring polypeptide capable of converting compound (2) to compound (1) with improved properties relative to the activity of the polypeptide of SEQ ID NO: 2, wherein the non-naturally occurring polypeptide comprises an amino acid sequence having at least 80%, 85%, 86%, 87%, 88%,

89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identity to SEQ ID NO: 2, and further comprises a set of amino acid residue differences as compared to SEQ ID NO:2, wherein the amino acid differences are based on locations or regions in the structure of reference polypeptide (e.g., SEQ ID NO: 2) and/or the associated function properties. Accordingly, referring to Table 3, a non-naturally occurring or engineered ketoreductase polypeptide of the present disclosure can include an amino acid substitution at a particular residue at a location in the structure of the reference polypeptide as identified in Table 3. Exemplary substitutions at each of the relevant locations include those identified in Table 2.

TABLE 3: Structural locations useful for engineered ketoreductase polypeptides

Position	Structural location	Associated functional properties
X21	Surface	Thermostability
X25	Surface	Solvent stability
X40	NADPH-Binding Site	Tight binding of NADPH to enzyme
X64	NADPH-Binding Site	Interacts with NADPH Adenine ring
X93	Second sphere active site	Thermostability
X94	Second sphere active site	Thermostability
X95	Second sphere active site	Activity
X96	Second sphere active site	Activity
X99	Dimer-tetramer interface	Thermostability/Solvent stability
X108	Dimer-tetramer interface	Thermostability/Solvent stability
X117	Core	Thermostability/Solvent stability
X127	Second sphere active site	Thermostability/Solvent stability
X147	Dimer-tetramer interface	Thermostability/Solvent stability
X148	Dimer-tetramer interface	Thermostability/Solvent stability
X150	Active site	Activity
X152	Second sphere active site	Activity
X153	Second sphere active site	Activity
X155	Position interacting with 95	Activity
X190	Active site	Activity
X195	Second sphere active site	Thermostability
X196	Active site	Activity
X201	Active site	Activity
X202	Flexible loop	Activity
X203	Flexible loop	Activity
X204	Flexible loop	Activity
X205	Flexible loop	Activity
X206	Flexible loop	Activity
X207	Flexible loop	Activity
X211	Second sphere active site	Activity
X221	Surface	Thermostability/Solvent stability
X223	Core	Thermostability/Solvent stability
X226	Dimer-tetramer interface	Thermostability/Solvent stability

[0113] In some embodiments, the present disclosure provides a non-naturally occurring polypeptide capable of converting compound (2) to compound (1) with at least 2-fold, at least 10-fold, at least 25-fold, at least 40-fold, or at least 60-fold increased activity relative to the activity of the polypeptide of SEQ ID NO: 2, which comprises an amino acid sequence having at least 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identity to SEQ ID NO: 2 and at least the following features: residue at position corresponding to **X40** is R; residue at position corresponding to **X153** is I, or L; residue at position corresponding to **X190** is A or P; residue at position corresponding to **X196** is T; residue at position corresponding to **X199** is F, or W; and residue at position corresponding to **X206** is I. In some embodiments, the amino acid sequence further comprises at least one feature or group of features selected from: (a) residue at position **X93** is A and residue at position **X94** is T; (b) residue at position **X93** is A and residue at position **X94** is S; (c) residue at position **X93** is A and residue at position **X94** is S; (d) residue at position **X93** is I and residue at position **X94** is S; (e) residue at position **X203** is G; (f) residue at position **X202** is G and residue at position **X203** is G; or (f) residue at position **X201** is A, residue at position **X202** is G, and residue at position **X203** is G.

[0114] In some embodiments, any of the non-naturally occurring polypeptides of the present disclosure capable of converting compound (2) to compound (1) with at least 2-fold, at least 10-fold, at least 25-fold, at least 40-fold, or at least 60-fold increased activity relative to the activity of the polypeptide of SEQ ID NO: 2 and an amino acid sequence having at least 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identity to SEQ ID NO: 2, can further one or more features selected from: residue at position corresponding to **X21** is R or F; residue at position corresponding to **X25** is R, T, or N; residue at position corresponding to **X40** is R; residue at position corresponding to **X64** is V; residue at position corresponding to **X93** is A; residue at position corresponding to **X94** is T, S, or P; residue at position corresponding to **X95** is V, or M; residue at position corresponding to **X96** is V, G, A, N, S, P, or T; residue at position corresponding to **X99** is L; residue at position corresponding to **X108** is H; residue at position corresponding to **X117** is A, or G; residue at position corresponding to **X127** is K, or Q; residue at position corresponding to **X147** is M, or I; residue at position corresponding to **X148** is I; residue at position corresponding to **X150** is H, or A; residue at position corresponding to **X152** is N, or F; residue at position corresponding to **X153** is I, or L; residue at position corresponding to **X155** is C; residue at position corresponding to **X190** is A; residue at position corresponding to **X195** is M; residue at position corresponding to **X196** is T, A, S, C, or N; residue at position corresponding to **X199** is F, or W; residue at position corresponding to **X201** is I, L, or A; residue at position corresponding to **X202** is L, N, V, or G; residue at position corresponding to **X203** is G; residue at position corresponding to **X204** is V, or A; residue at position corresponding to **X205** is V; residue at position corresponding to **X206** is I; residue at position corresponding to **X207** is T, C, I, or N; residue at position corresponding to **X211** is K; residue at position corresponding to **X221** is D; residue at position corresponding to **X223** is I; residue at position corresponding to **X226** is V.

[0115] In some embodiments, the present disclosure provides a non-naturally occurring polypeptide capable of converting compound (2) to compound (1) with at least 2-fold, at least 10-fold, at least 25-fold, at least 40-fold, or at least 60-fold increased activity relative to the activity of the polypeptide of SEQ ID NO: 2, which comprises an amino acid sequence having at least 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identity to SEQ ID NO: 2, and at least the following features: residue at position corresponding to **X40** is R; residue at position corresponding to **X153** is I, or L; residue at position corresponding to **X190** is A or P; residue at position corresponding to **X196** is T; residue at position corresponding to **X199** is F, or W; and residue at position corresponding to **X206** is I, wherein the amino acid sequence further comprises at least one of the following sets of features (a) through (h):

[0116] (a) residue at position corresponding to **X40** is R; residue at position corresponding to **X93** is A; residue at position corresponding to **X94** is T; residue at position corresponding to **X153** is I; residue at position corresponding to **X190** is A; residue at position corresponding to **X196** is T; residue at position corresponding to **X199** is F; residue at position corresponding to **X203** is G; and residue at position corresponding to **X206** is I; or

[0117] (b) residue at position corresponding to **X40** is R; residue at position corresponding to **X93** is A; residue at position corresponding to **X94** is S; residue at position corresponding to **X96** is N; residue at position corresponding to **X153** is L; residue at position corresponding to **X190** is A; residue at position corresponding to **X196** is T; residue at position corresponding to **X199** is F; residue at position corresponding to **X202** is G; residue at position corresponding to **X203** is G; and residue at position corresponding to **X206** is I; or

[0118] (c) residue at position corresponding to **X40** is R; residue at position corresponding to **X93** is A; residue at position corresponding to **X94** is S; residue at position corresponding to **X96** is A, P, or N; residue at position corresponding to **X153** is L; residue at position corresponding to **X190** is A; residue at position corresponding to **X196** is T; residue at position corresponding to **X199** is F; residue at position corresponding to **X202** is G; residue at position corresponding to **X203** is G; and residue at position corresponding to **X206** is I; and, optionally further comprises: residue at position corresponding to **X96** is N; and residue at position corresponding to **X201** is I, or L; or

[0119] (d) residue at position corresponding to **X40** is R; residue at position corresponding to **X93** is A; residue at position corresponding to **X96** is P; residue at position corresponding to **X153** is L; residue at position corresponding to **X190** is A; residue at position corresponding to **X196** is T; residue at position corresponding to **X199** is F; residue at position corresponding to **X202** is G; residue at position corresponding to **X203** is G; and residue at position corresponding

to **X206** is I; and optionally further comprises: residue at position corresponding to **X21** is F; residue at position corresponding to **X25** is R; residue at position corresponding to **X94** is S; residue at position corresponding to **X96** is P; residue at position corresponding to **X108** is K; residue at position corresponding to **X127** is K; residue at position corresponding to **X147** is M; residue at position corresponding to **X153** is L; residue at position corresponding to **X205** is V; residue at position corresponding to **X211** is K; and residue at position corresponding to **X221** is D; or

[0120] (e) residue at position corresponding to **X21** is F; residue at position corresponding to **X25** is R; residue at position corresponding to **X40** is R; residue at position corresponding to **X93** is A; residue at position corresponding to **X94** is S; residue at position corresponding to **X96** is P; residue at position corresponding to **X108** is K; residue at position corresponding to **X127** is K; residue at position corresponding to **X147** is M; residue at position corresponding to **X153** is L; residue at position corresponding to **X190** is A; residue at position corresponding to **X196** is T; residue at position corresponding to **X199** is F; residue at position corresponding to **X202** is G; residue at position corresponding to **X203** is G; residue at position corresponding to **X206** is I; and residue at position corresponding to **X221** is D; or

[0121] (f) residue at position corresponding to **X40** is R; residue at position corresponding to **X93** is I; residue at position corresponding to **X94** is S; residue at position corresponding to **X96** is P; residue at position corresponding to **X153** is L; residue at position corresponding to **X190** is A; residue at position corresponding to **X196** is T; residue at position corresponding to **X199** is F, or W; residue at position corresponding to **X201** is A; residue at position corresponding to **X202** is G; residue at position corresponding to **X203** is G; residue at position corresponding to **X206** is I; and residue at position corresponding to **X211** is K; and optionally further comprises: residue at position corresponding to **X64** is V; residue at position corresponding to **X99** is L; residue at position corresponding to **X108** is H; residue at position corresponding to **X117** is A, or G; residue at position corresponding to **X147** is I, or M; residue at position corresponding to **X195** is M; residue at position corresponding to **X205** is V; residue at position corresponding to **X223** is I; and residue at position corresponding to **X226** is V; or

[0122] (g) residue at position corresponding to **X40** is R; residue at position corresponding to **X93** is I; residue at position corresponding to **X94** is S; residue at position corresponding to **X96** is P; residue at position corresponding to **X99** is L; residue at position corresponding to **X108** is H; residue at position corresponding to **X117** is G; residue at position corresponding to **X147** is I; residue at position corresponding to **X153** is L; residue at position corresponding to **X190** is A; residue at position corresponding to **X196** is T; residue at position corresponding to **X199** is F; residue at position corresponding to **X201** is A; residue at position corresponding to **X202** is G; residue at position corresponding to **X203** is G; residue at position corresponding to **X206** is I; residue at position corresponding to **X211** is K; and residue at position corresponding to **X226** is V; or

[0123] (h) residue at position corresponding to **X40** is R; residue at position corresponding to **X64** is V; residue at position corresponding to **X93** is I; residue at position corresponding to **X94** is S; residue at position corresponding to **X96** is P; residue at position corresponding to **X99** is L; residue at position corresponding to **X108** is H; residue at position corresponding to **X117** is G; residue at position corresponding to **X147** is I; residue at position corresponding to **X153** is L; residue at position corresponding to **X190** is A; residue at position corresponding to **X196** is T; residue at position corresponding to **X199** is W; residue at position corresponding to **X201** is A; residue at position corresponding to **X202** is G; residue at position corresponding to **X203** is G; residue at position corresponding to **X206** is I; residue at position corresponding to **X211** is K; and residue at position corresponding to **X226** is V.

[0124] In some embodiments, the non-naturally occurring polypeptides of the present disclosure can comprise a sequence having one or more amino acid residue differences as compared to SEQ ID NO: 2 at residue positions **affecting activity** for conversion of compound (2) to compound (1), which positions include the following: **X21; X25; X64; X93; X94; X95; X96; X99; X108; X117; X127; X147; X148; X150; X152; X153; X155; X163; X190; X195; X196; X199; X201; X202; X203; X204; X205; X206; X207; X211; X221; X223; and X226**. In some embodiments, the specific amino acid differences **resulting in increased activity** for conversion of compound (2) to compound (1) relative to the reference polypeptide of SEQ ID NO: 2 can be selected from the following: residue at position corresponding to **X21** is F or R; residue at position corresponding to **X25** is N, R, or T; residue at position corresponding to **X64** is V; residue at position corresponding to **X93** is A or T; residue at position corresponding to **X94** is P, S, or T; residue at position corresponding to **X95** is M or V; residue at position corresponding to **X96** is A, N, G, P, S, T, or V; residue at position corresponding to **X99** is L; residue at position corresponding to **X108** is D, H, or K; residue at position corresponding to **X117** is A or G; residue at position corresponding to **X127** is K or Q; residue at position corresponding to **X147** is I or M; residue at position corresponding to **X148** is I; residue at position corresponding to **X150** is H; residue at position corresponding to **X152** is N or F; residue at position corresponding to **X153** is I or L; residue at position corresponding to **X155** is C; residue at position corresponding to **X163** is I; residue at position corresponding to **X190** is A; residue at position corresponding to **X195** is M; residue at position corresponding to **X196** is A, C, N, S, or T; residue at position corresponding to **X199** is F or W; residue at position corresponding to **X201** is A, I, or L; residue at position corresponding to **X202** is G, L, N, or V; residue at position corresponding to **X203** is G; residue at position corresponding to **X204** is A or V; residue at position corresponding to **X205** is V; residue at position corresponding to **X206** is I; residue at position corresponding to **X207** is C, I, N, or T; residue at position corresponding to **X211** is K; residue at position corresponding to **X221** is D;

residue at position corresponding to **X223** is I; and residue at position corresponding to **X226** is V.

[0125] In some embodiments, the non-naturally occurring polypeptides capable of converting compound **(2)** to compound **(1)** can have increased thermostability as compared to the polypeptide of SEQ ID NO: 2 (or another reference polypeptide, e.g., SEQ ID NO: 80 or 100). Thermostability can be determined by preincubating the polypeptide at a defined temperature and time, e.g., 4°C - 46°C for 18-24 hours, followed by measuring the % residual activity using a defined assay. Exemplary preincubation conditions include preincubation at 30°C for 18 h, or 40°C for 24 h. Accordingly, in some embodiments, the non-naturally occurring polypeptides of the present disclosure can comprise an amino acid sequence having one or more residue differences as compared to SEQ ID NO: 2 at residue positions affecting **thermostability**, which positions include the following: **X21**; **X93**; **X94**; **X117**; **X127**; **X147**; **X195**; and **X199**. In some embodiments, specific amino acid differences resulting in **increased thermostability** relative to the reference polypeptide of SEQ ID NO: 2 can be selected from the following substitutions: residue at position corresponding to **X21** is F; residue at position corresponding to **X93** is T; residue at position corresponding to **X94** is A; residue at position corresponding to **X117** is G or A; residue at position corresponding to **X127** is K; residue at position corresponding to **X147** is I; residue at position corresponding to **X195** is M; and residue at position corresponding to **X199** is W.

[0126] In some embodiments, the present disclosure provides non-naturally occurring polypeptides capable of converting compound **(2)** to compound **(1)** and having at least 1.5-fold, 2.5-fold, 5-fold, 7.5-fold or more increased thermostability following 18 h preincubation at 40°C as compared to the polypeptide of SEQ ID NO: 80 or SEQ ID NO: 100, which comprises an amino acid sequence having at least 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identity to SEQ ID NO: 80 or SEQ ID NO: 100, and at least one of the following substitutions: residue at position corresponding to **X21** is F; residue at position corresponding to **X93** is T; residue at position corresponding to **X94** is A; residue at position corresponding to **X117** is G or A; residue at position corresponding to **X127** is K; residue at position corresponding to **X147** is I; residue at position corresponding to **X195** is M; and residue at position corresponding to **X199** is W.

[0127] In some embodiments, the present disclosure provides non-naturally occurring polypeptides capable of converting compound **(2)** to compound **(1)** and having at least 1.5-fold, 2.5-fold, 5-fold, 7.5-fold or more increased thermostability following 18 h preincubation at 40°C as compared to the polypeptide of SEQ ID NO: 80 or SEQ ID NO: 100, which comprises an amino acid sequence having at least 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identity to SEQ ID NO: 80 or SEQ ID NO: 100, and the set of amino acid residue differences of any one of the non-naturally occurring polypeptides of SEQ ID NOs: 134, 136, 140, 142, 144, 152, 154, 156, 158, 160, 162, 164, or 166.

[0128] In some embodiments, the non-naturally occurring polypeptides of the present disclosure can comprise an amino acid sequence having residue differences as compared to SEQ ID NO: 2 at residue positions affecting **solvent stability**, which positions include the following: **X25**; **X147**; and **X221**. In some embodiments, specific amino acid differences resulting in **increased solvent stability** relative to the reference polypeptide of SEQ ID NO: 2 (e.g., increased activity relative to SEQ ID NO: 2 in up to 65% isopropanol) can be selected from the following substitutions: residue at position corresponding to **X25** is R; residue at position corresponding to **X147** is M; and residue at position corresponding to **X221** is D.

[0129] In some embodiments, the non-naturally occurring polypeptides of the present disclosure can comprise an amino acid sequence having residue differences as compared to SEQ ID NO: 2 at residue positions affecting **cofactor binding**, which positions include **X40**. In some embodiments, specific amino acid differences affecting cofactor binding can be selected from the following: residue at position corresponding to **X40** is R.

[0130] As will be apparent to the skilled artisan, various combinations of residue differences as compared to SEQ ID NO: 2 at residue positions affecting enzymatic activity, thermostability, solvent stability, and cofactor binding can be made to form the polypeptides of the present disclosure.

[0131] In addition to the residue positions specified above, any of the non-naturally occurring ketoreductase polypeptides disclosed herein can further comprise other residue differences relative to SEQ ID NO: 2 at **other** residue positions. Residue differences at these other residue positions provide for additional variations in the amino acid sequence without adversely affecting the ability of the polypeptide to carry out the conversion of compound **(2)** to compound **(1)**. In some embodiments, the polypeptides can have additionally 1-2, 1-3, 1-4, 1-5, 1-6, 1-7, 1-8, 1-9, 1-10, 1-11, 1-12, 1-14, 1-15, 1-16, 1-18, 1-20, 1-22, 1-24, 1-26, 1-30, 1-35, 1-40 residue differences at other amino acid residue positions as compared to the reference sequence. In some embodiments, the number of differences can be 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 15, 16, 18, 20, 22, 24, 26, 30, 35, and 40 residue differences at other residue positions. The residue difference at these other positions can include conservative changes or non-conservative changes. In some embodiments, the residue differences can comprise conservative substitutions and non-conservative substitutions as compared to the wild-type ketoreductase of SEQ ID NO: 2.

[0132] Amino acid residue differences at other positions relative to SEQ ID NO: 2 or the wild-type *L. kefir* ketoreductase sequence (Genbank acc. No. AAP94029.1; GI: 33112056) and the affect of these differences on enzyme function are provide by e.g., engineered ketoreductase polypeptides in the following patent publications: US 20080318295A1; US

20090093031A1; US 20090155863A1; US 20090162909A1; US 20090191605A1; US 20100055751A1; WO/2010/025238A2; WO/2010/025287A2; and US 20100062499A1. Accordingly, in some embodiments, one or more of the amino acid differences provided in the engineered ketoreductase polypeptides of these publications could also be introduced into a non-naturally occurring polypeptide of the present disclosure.

[0133] In some embodiments, the present disclosure provides a non-naturally occurring polypeptide capable of converting compound (2) to compound (1) with at least 2-fold, at least 10-fold, at least 25-fold, at least 40-fold, or at least 60-fold increased activity relative to the activity of the polypeptide of SEQ ID NO: 2, which comprises an amino acid sequence having at least 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identity to SEQ ID NO: 2, with the proviso that the amino acid sequence of any one or more of the ketoreductase polypeptides disclosed in any one or more of the following patent publications are excluded: US 20080318295A1; US 20090093031A1; US 20090155863A1; US 20090162909A1; US 20090191605A1; US 20100055751A1; WO/2010/025238A2; WO/2010/025287A2; US 20100062499A1; and WO 2008/151324A1.

[0134] In some embodiments, the polypeptides can comprise deletions of the engineered ketoreductase polypeptides described herein. Thus, for each and every embodiment of the polypeptides of the disclosure, the deletions can comprise one or more amino acids, 2 or more amino acids, 3 or more amino acids, 4 or more amino acids, 5 or more amino acids, 6 or more amino acids, 8 or more amino acids, 10 or more amino acids, 15 or more amino acids, or 20 or more amino acids, up to 10% of the total number of amino acids, up to 10% of the total number of amino acids, up to 20% of the total number of amino acids of the polypeptides, as long as the functional activity of the polypeptide with respect to the conversion of compound (2) to compound (1) is present. In some embodiments, the deletions can comprise, 1-2, 1-3, 1-4, 1-5, 1-6, 1-7, 1-8, 1-9, 1-10, 1-11, 1-12, 1-14, 1-15, 1-16, 1-18, 1-20, 1-22, 1-24, 1-26, 1-30, 1-35, or 1-40 amino acid residues. In some embodiments, the number of deletions can be 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 15, 16, 18, 20, 22, 24, 26, 30, 35, or 40 amino acids. In some embodiments, the deletions can comprise deletions of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 18, or 20 amino acid residues.

[0135] In some embodiments, the polypeptides can comprise fragments of the engineered polypeptides described herein. In some embodiments, the fragments can have about 80%, 90%, 95%, 98%, and 99% of the full-length polypeptide, e.g., the polypeptide of SEQ ID NO:2, as long as the functional activity of the polypeptide with respect to the conversion of compound (2) to compound (1) is present.

[0136] In some embodiments, the polypeptides of the disclosure can be in the form of fusion polypeptides in which the engineered polypeptides are fused to other polypeptides, such as, by way of example and not limitation, antibody tags (e.g., myc epitope), purifications sequences (e.g., His tags for binding to metals), and cell localization signals (e.g., secretion signals). Thus, the engineered polypeptides described herein can be used with or without fusions to other polypeptides.

[0137] As will be understood by the skilled artisan, the polypeptides described herein are not restricted to the genetically encoded amino acids. In addition to the genetically encoded amino acids, the polypeptides described herein may be comprised, either in whole or in part, of naturally-occurring and/or synthetic non-encoded amino acids. Certain commonly encountered non-encoded amino acids of which the polypeptides described herein may be comprised include, but are not limited to: the D-enantiomers of the genetically-encoded amino acids; 2,3-diaminopropionic acid (Dpr); α -aminoisobutyric acid (Aib); ϵ -aminohexanoic acid (Aha); δ -aminovaleric acid (Ava); N-methylglycine or sarcosine (MeGly or Sar); ornithine (Orn); citrulline (Cit); t-butylalanine (Bua); t-butylglycine (Bug); N-methylisoleucine (Melle); phenylglycine (Phg); cyclohexylalanine (Cha); norleucine (Nle); naphthylalanine (Nal); 2-chlorophenylalanine (Ocf); 3-chlorophenylalanine (Mcf); 4-chlorophenylalanine (Pcf); 2-fluorophenylalanine (Off); 3-fluorophenylalanine (Mff); 4-fluorophenylalanine (Pff); 2-bromophenylalanine (Obf); 3-bromophenylalanine (Mbf); 4-bromophenylalanine (Pbf); 2-methylphenylalanine (Omf); 3-methylphenylalanine (Mmf); 4-methylphenylalanine (Pmf); 2-nitrophenylalanine (Onf); 3-nitrophenylalanine (Mnf); 4-nitrophenylalanine (Pnf); 2-cyanophenylalanine (Ocf); 3-cyanophenylalanine (Mcf); 4-cyanophenylalanine (Pcf); 2-trifluoromethylphenylalanine (Otf); 3-trifluoromethylphenylalanine (Mtf); 4-trifluoromethylphenylalanine (Ptf); 4-aminophenylalanine (Paf); 4-iodophenylalanine (Pif); 4-aminomethylphenylalanine (Pamf); 2,4-dichlorophenylalanine (Opef); 3,4-dichlorophenylalanine (Mpcf); 2,4-difluorophenylalanine (Opff); 3,4-difluorophenylalanine (Mpff); pyrid-2-ylalanine (2pAla); pyrid-3-ylalanine (3pAla); pyrid-4-ylalanine (4pAla); naphth-1-ylalanine (1nAla); naphth-2-ylalanine (2nAla); thiazolylalanine (taAla); benzothienylalanine (bAla); thienylalanine (tAla); furylalanine (fAla); homophenylalanine (hPhe); homotyrosine (hTyr); homotryptophan (hTrp); pentafluorophenylalanine (5ff); styrylalanine (sAla); authrylalanine (aAla); 3,3-diphenylalanine (Dfa); 3-amino-5-phenylpentanoic acid (Afp); penicillamine (Pen); 1,2,3,4-tetrahydroisoquinoline-3-carboxylic acid (Tic); β -2-thienylalanine (Thi); methionine sulfoxide (Mso); N(w)-nitroarginine (nArg); homolysine (hLys); phosphonomethylphenylalanine (pmPhe); phosphoserine (pSer); phosphothreonine (pThr); homoaspartic acid (hAsp); homoglutamic acid (hGlu); 1-aminocyclopent-(2 or 3)-ene-4 carboxylic acid; pipercolic acid (PA), azetidine-3-carboxylic acid (ACA); 1-aminocyclopentane-3-carboxylic acid; allylglycine (aOly); propargylglycine (pgGly); homoalanine (hAla); norvaline (nVal); homoleucine (hLeu), homovaline (hVal); homoisoleucine (hIle); homoarginine (hArg); N-acetyl lysine (AcLys); 2,4-diaminobutyric acid (Dbu); 2,3-diaminobutyric acid (Dab); N-methylvaline (MeVal); homocysteine (hCys); homoserine (hSer); hydroxyproline (Hyp) and homoproline (hPro). Additional non-encoded amino

acids of which the polypeptides described herein may be comprised will be apparent to those of skill in the art (see, e.g., the various amino acids provided in Fasman, 1989, CRC Practical Handbook of Biochemistry and Molecular Biology, CRC Press, Boca Raton, FL, at pp. 3-70 and the references cited therein). These amino acids may be in either the L- or D-configuration.

[0138] Those of skill in the art will recognize that amino acids or residues bearing side chain protecting groups may also comprise the polypeptides described herein. Non-limiting examples of such protected amino acids, which in this case belong to the aromatic category, include (protecting groups listed in parentheses), but are not limited to: Arg(tos), Cys(methylbenzyl), Cys (nitropyridinesulfonyl), Glu(δ -benzylester), Gln(xanthyl), Asn(N- δ -xanthyl), His(bom), His(benzyl), His(tos), Lys(fmoc), Lys(tos), Ser(O-benzyl), Thr (O-benzyl) and Tyr(O-benzyl).

[0139] Non-encoding amino acids that are conformationally constrained of which the polypeptides described herein may be composed include, but are not limited to, N-methyl amino acids (L-configuration); 1-aminocyclopent-(2 or 3)-ene-4-carboxylic acid; pipercolic acid; azetidine-3-carboxylic acid; homoproline (hPro); and 1-aminocyclopentane-3-carboxylic acid.

[0140] In some embodiments, the polypeptide described herein can be provided in the form of kits. The enzymes in the kits may be present individually or as a plurality of enzymes. The kits can further include reagents for carrying out the enzymatic reactions, substrates for assessing the activity of enzymes, as well as reagents for detecting the products. The kits can also include reagent dispensers and instructions for use of the kits.

[0141] In some embodiments, the polypeptides can be provided on a substrate. In some embodiments, the polypeptides can be provided in the form of an array in which the polypeptides are arranged in positionally distinct locations. The array can be used to test a variety of aryl alkyl sulfides for conversion by the polypeptides. "Substrate," "support," "solid support," "solid carrier," or "resin" in the context of arrays refer to any solid phase material. Substrate also encompasses terms such as "solid phase," "surface," and/or "membrane." A solid support can be composed of organic polymers such as polystyrene, polyethylene, polypropylene, polyfluoroethylene, polyethyleneoxy, and polyacrylamide, as well as co-polymers and grafts thereof. A solid support can also be inorganic, such as glass, silica, controlled pore glass (CPG), reverse phase silica or metal, such as gold or platinum. The configuration of a substrate can be in the form of beads, spheres, particles, granules, a gel, a membrane or a surface. Surfaces can be planar, substantially planar, or non-planar. Solid supports can be porous or non-porous, and can have swelling or non-swelling characteristics. A solid support can be configured in the form of a well, depression, or other container, vessel, feature, or location. A plurality of supports can be configured on an array at various locations, addressable for robotic delivery of reagents, or by detection methods and/or instruments.

[0142] In certain embodiments, the kits of the present disclosure include arrays comprising a plurality of different engineered ketoreductase polypeptides at different addressable position, wherein the different polypeptides are different variants of a reference sequence each having at least one different improved enzyme property. Such arrays comprising a plurality of engineered polypeptides and methods of their use are described in, e.g., WO2009/008908A2.

4.4 Ketoreductase Polynucleotides, Expression Vectors, and Host Cells

[0143] In another aspect, the present disclosure provides polynucleotides encoding the non-naturally occurring or engineered polypeptides described herein. These polynucleotides may be operatively linked to one or more heterologous regulatory sequences that control gene expression to create a recombinant polynucleotide capable of expressing the ketoreductase polypeptide. Expression constructs containing a heterologous polynucleotide encoding the engineered ketoreductase polypeptide can be introduced into appropriate host cells to express the corresponding polypeptide.

[0144] Because of the knowledge of the codons corresponding to the various amino acids, availability of a protein sequence provides a description of all the polynucleotides capable of encoding the subject. Thus, having identified a particular amino acid sequence, those skilled in the art could make any number of different nucleic acids by simply modifying the sequence of one or more codons in a way which does not change the amino acid sequence of the protein. In this regard, the present disclosure specifically contemplates each and every possible variation of polynucleotides that could be made by selecting combinations based on the possible codon choices, and all such variations are to be considered specifically disclosed for any polypeptide disclosed herein, including the amino acid sequences presented in Table 2.

[0145] In some embodiments, the polynucleotides can be selected and/or engineered to comprise codons that are preferably selected to fit the host cell in which the protein is being produced. For example, preferred codons used in bacteria are used to express the gene in bacteria; preferred codons used in yeast are used for expression in yeast; and preferred codons used in mammals are used for expression in mammalian cells. Since not all codons need to be replaced to optimize the codon usage of the ketoreductases (e.g., because the natural sequence can have preferred codons and because use of preferred codons may not be required for all amino acid residues), codon optimized polynucleotides encoding the ketoreductase polypeptides may contain preferred codons at about 40%, 50%, 60%, 70%, 80%, or greater than 90% of codon positions of the full length coding region.

[0146] In some embodiments, the polynucleotide encodes a non-naturally occurring polypeptide capable of converting compound (2) to compound (1) with at least 2-fold, at least 10-fold, at least 25-fold, at least 40-fold, or at least 60-fold increased activity relative to the activity of the polypeptide of SEQ ID NO: 2, and comprises an amino acid sequence having at least 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identity to a reference amino acid sequence selected from any one of SEQ ID NO: 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, or 168.

[0147] In some embodiments, the polynucleotide encodes a non-naturally occurring polypeptide capable of converting compound (2) to compound (1) with at least 2-fold, at least 10-fold, at least 25-fold, at least 40-fold, or at least 60-fold increased activity relative to the activity of the polypeptide of SEQ ID NO: 2, and comprises an amino acid sequence having at least 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identity to a reference amino acid sequence selected from any one of SEQ ID NO: 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, or 168, with the proviso that the amino acid sequence comprises any one of the set of residue differences as compared to SEQ ID NO: 2 contained in any one of the polypeptide sequences of SEQ ID NO: 4 to SEQ ID NO: 168 listed in Table 2.

[0148] In some embodiments, the polynucleotides encoding the polypeptides are selected from SEQ ID NO: 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139, 141, 143, 145, 147, 149, 151, 153, 155, 157, 159, 161, 163, 165, and 167.

[0149] In some embodiments, the polynucleotides are capable of hybridizing under highly stringent conditions to a polynucleotide comprising SEQ ID NO: 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139, 141, 143, 145, 147, 149, 151, 153, 155, 157, 159, 161, 163, 165, or 167, or a complement thereof, where the highly stringently hybridizing polynucleotides encode a non-naturally occurring polypeptide capable of converting compound (2) to compound (1) with at least 2-fold, at least 10-fold, at least 25-fold, at least 40-fold, or at least 60-fold increased activity relative to the activity of the polypeptide of SEQ ID NO: 2.

[0150] In some embodiments, the polynucleotides encode the polypeptides described herein but have about 80% or more sequence identity, about 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% or more sequence identity at the nucleotide level to a reference polynucleotide encoding the engineered ketoreductase polypeptides described herein. In some embodiments, the reference polynucleotide is selected from SEQ ID NO: 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139, 141, 143, 145, 147, 149, 151, 153, 155, 157, 159, 161, 163, 165, and 167.

[0151] An isolated polynucleotide encoding a non-naturally occurring polypeptide of the disclosure may be manipulated in a variety of ways to provide for expression of the polypeptide. In some embodiments, the polynucleotides encoding the polypeptides can be provided as expression vectors where one or more control sequences is present to regulate the expression of the polynucleotides. Manipulation of the isolated polynucleotide prior to its insertion into a vector may be desirable or necessary depending on the expression vector. The techniques for modifying polynucleotides and nucleic acid sequences utilizing recombinant DNA methods are well known in the art. Guidance is provided in Sambrook et al., 2001, *Molecular Cloning: A Laboratory Manual*, 3rd Ed., Cold Spring Harbor Laboratory Press; and *Current Protocols in Molecular Biology*, Ausubel, F. ed., Greene Pub. Associates, 1998, updates to 2006.

[0152] In some embodiments, the control sequences include among others, promoters, leader sequence, polyadenylation sequence, propeptide sequence, signal peptide sequence, and transcription terminator. Suitable promoters can be selected based on the host cells used. Exemplary bacterial promoters include *E. coli* lac operon, *E. coli* trp operon, bacteriophage ϕ 1, *Streptomyces coelicolor* agarase gene (*dagA*), *Bacillus subtilis* levansucrase gene (*sacB*), *Bacillus licheniformis* alpha-amylase gene (*amyL*), beta-lactamase gene, and *tac* promoter; exemplary promoters for filamentous fungal host cells, include promoters obtained from the genes for *Aspergillus oryzae* TAKA amylase, *Rhizomucor miehei* aspartic proteinase, *Aspergillus niger* neutral alpha-amylase, *Aspergillus niger* acid stable alpha-amylase, *Aspergillus niger* or *Aspergillus awamori* glucoamylase (*glaA*), *Rhizomucor miehei* lipase, *Aspergillus oryzae* alkaline protease, *Aspergillus oryzae* triose phosphate isomerase, *Aspergillus nidulans* acetamidase, and *Fusarium oxysporum* trypsin-like protease, and mutant, truncated, and hybrid promoters thereof, and exemplary yeast cell promoters can be from the genes for *Saccharomyces cerevisiae* enolase (*ENO-1*), *Saccharomyces cerevisiae* galactokinase (*GAL1*), *Saccharomyces cerevisiae* alcohol dehydrogenase/glyceraldehyde-3-phosphate dehydrogenase (*ADH2/GAP*), and *Saccharomyces cerevisiae* 3-phosphoglycerate kinase.

[0153] In some embodiments, the control sequence may also be a signal peptide coding region that codes for an amino acid sequence linked to the amino terminus of a polypeptide and directs the encoded polypeptide into the cell's secretory pathway. The signal sequence typically depends on the type of host cells being used to express the polypeptide. Effective signal peptide coding regions for bacterial host cells are the signal peptide coding regions obtained from the genes for *Bacillus* NC1B 11837 maltogenic amylase, *Bacillus stearothermophilus* alpha-amylase, *Bacillus licheniformis* subtilisin, *Bacillus licheniformis* beta-lactamase, *Bacillus stearothermophilus* neutral proteases (nprT, nprS, nprM), and *Bacillus subtilis* prsA. Exemplary signal peptide coding regions for filamentous fungal host cells can be the signal peptide coding regions obtained from the genes for *Aspergillus oryzae* TAKA amylase, *Aspergillus niger* neutral amylase, *Aspergillus niger* glucoamylase, *Rhizomucor miehei* aspartic proteinase, *Humicola insolens* cellulase, and *Humicola lanuginosa* lipase. Useful signal peptides for yeast host cells can be from the genes for *Saccharomyces cerevisiae* alpha-factor and *Saccharomyces cerevisiae* invertase.

[0154] Other control sequences, such as leader sequence, polyadenylation sequence, and transcription terminator sequences can use those available in the art (see Sambrook, supra, and Current Protocols in Molecular Biology, supra).

[0155] In another aspect, the present disclosure is also directed to a recombinant expression vector comprising a polynucleotide encoding an engineered ketoreductase polypeptide or a variant thereof, and one or more expression regulating regions such as a promoter and a terminator, a replication origin, etc., depending on the type of hosts into which they are to be introduced. The recombinant expression vector may be any vector (e.g., a plasmid or virus), which can be conveniently subjected to recombinant DNA procedures and can bring about the expression of the polynucleotide sequence. 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 vectors may be linear or closed circular plasmids.

[0156] The expression 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 which, when introduced into the host cell, is integrated into the genome and replicated together with the chromosome(s) into which it has been integrated. The expression vector preferably contains one or more selectable markers, which permit easy selection of transformed cells. A selectable marker is a gene the product of which provides for biocide or viral resistance, resistance to heavy metals, prototrophy to auxotrophs, resistance to chemical agents (e.g., antibiotics) and the like.

[0157] In another aspect, the present disclosure provides a host cell comprising a polynucleotide encoding an engineered ketoreductase polypeptide of the present disclosure, the polynucleotide being operatively linked to one or more control sequences for expression of the ketoreductase polypeptide in the host cell. Host cells for use in expressing the ketoreductase polypeptides encoded by the expression vectors of the present invention are well known in the art and include but are not limited to, bacterial cells, such as *E. coli*, *Lactobacillus*, *Streptomyces* and *Salmonella typhimurium* cells; fungal cells, such as yeast cells; insect cells such as *Drosophila* S2 and *Spodoptera* Sf9 cells; animal cells such as CHO, COS, BHK, 293, and Bowes melanoma cells; and plant cells. Exemplary host cells are *Escherichia coli* BL21 and W3110.

[0158] Appropriate culture mediums and growth conditions for the above-described host cells are well known in the art. Polynucleotides for expression of the ketoreductase may be introduced into host cells by various methods known in the art (e.g., electroporation, biolistic particle bombardment, liposome mediated transfection, calcium chloride transfection, and protoplast fusion).

[0159] In the embodiments herein, the non-naturally occurring or engineered ketoreductase polypeptides and nucleotides encoding such polypeptides can be prepared using methods commonly used by those skilled in the art. As noted above, the naturally-occurring amino acid sequence and corresponding polynucleotide encoding the ketoreductase enzyme of *Lactobacillus kefir*. In some embodiments, the parent polynucleotide sequence is codon optimized to enhance expression of the ketoreductase in a specified host cell.

[0160] The engineered ketoreductase polypeptides can be obtained by subjecting the polynucleotide encoding the naturally occurring ketoreductase to mutagenesis and/or directed evolution methods (see e.g., Stemmer, 1994, Proc Natl Acad Sci USA 91:10747-10751; PCT Publ. Nos. WO 95/22625, WO 97/0078, WO 97/35966, WO 98/27230, WO 00/42651, and WO 01/75767; U.S. Pat. Nos. 6,537,746, 6,117,679, 6,376,246, and 6,586,182; and U.S. Pat. Publ. Nos. 20080220990A1 and 20090312196A1).

[0161] Other directed evolution procedures that can be used include, among others, staggered extension process (StEP), in vitro recombination (Zhao et al., 1998, Nat. Biotechnol. 16:258-261), mutagenic PCR (Caldwell et al., 1994, PCR Methods Appl. 3:S136-S140), and cassette mutagenesis (Black et al., 1996, Proc Natl Acad Sci USA 93:3525-3529). Mutagenesis and directed evolution techniques useful for the purposes herein are also described in the following references: Ling, et al., 1997, Anal. Biochem. 254(2):157-78; Dale et al., 1996, Methods Mol. Biol. 57:369-74; Smith, 1985, Ann. Rev. Genet. 19:423-462; Botstein et al., 1985, Science 229:1193-1201; Carter, 1986, "Site-directed mutagenesis," Biochem. J. 237:1-7; Kramer et al., 1984, Cell 38:879-887; Wells et al., 1985, Gene 34:315-323; Minshull et al., 1999, Curr Opin Chem Biol 3:284-290; Christians et al., 1999, Nature Biotech 17:259-264; Crameri et al., 1998, Nature

391:288-291; Cramer et al., 1997, Nature Biotech 15:436-438; Zhang et al., 1997, Proc Natl Acad Sci USA 94:45-4-4509; Cramer et al., 1996, Nature Biotech 14:315-319; and Stemmer, 1994, Nature 370:389-391.

[0162] In some embodiments, the clones obtained following mutagenesis treatment are screened for non-naturally occurring ketoreductases having a desired enzyme property. Measuring ketoreductase enzyme activity from the expression libraries can be performed using the standard techniques, such as separation of the product (e.g., by HPLC) and detection of the product by measuring UV absorbance of the separated substrate and products and/or by detection using tandem mass spectroscopy (e.g., MS/MS). Clones containing a polynucleotide encoding the desired engineered polypeptides are then isolated, sequenced to identify the nucleotide sequence changes (if any), and used to express the enzyme in a host cell. Exemplary assays are provided below in Example 3.

[0163] Where the sequence of the polypeptide is known, the polynucleotides encoding the enzyme can be prepared by standard solid-phase methods, according to known synthetic methods, e.g., the classical phosphoramidite method described by Beaucage et al., 1981, Tet Lett 22:1859-69, or the method described by Matthes et al., 1984, EMBO J. 3:801-05. In some embodiments, fragments of up to about 100 bases can be individually synthesized, then joined (e.g., by enzymatic or chemical ligation methods, or polymerase mediated methods) to form any desired continuous sequence.

[0164] In some embodiments, the present disclosure also provides methods for preparing or manufacturing the non-naturally occurring polypeptides capable of converting compound (2) to compound (1), wherein the methods comprise: (a) culturing a host cell capable of expressing a polynucleotide encoding the non-naturally occurring polypeptide and (b) isolating the polypeptide from the host cell. The non-naturally occurring polypeptides can be expressed in appropriate cells (as described above), and isolated (or recovered) from the host cells and/or the culture medium using any one or more of the well known techniques used for protein purification, including, among others, lysozyme treatment, sonication, filtration, salting-out, ultra-centrifugation, and chromatography. Chromatographic techniques for isolation of the ketoreductase polypeptide include, among others, reverse phase chromatography high performance liquid chromatography, ion exchange chromatography, gel electrophoresis, and affinity chromatography.

[0165] In some embodiments, the non-naturally occurring polypeptide of the disclosure can be prepared and used in various isolated forms including but not limited to crude extracts (e.g., cell-free lysates), powders (e.g., shake-flask powders), lyophilizates, and substantially pure preparations (e.g., DSP powders), as further illustrated in the Examples below.

[0166] In some embodiments, the non-naturally occurring polypeptide of the disclosure can be prepared and used in purified form. Generally, conditions for purifying a particular enzyme will depend, in part, on factors such as net charge, hydrophobicity, hydrophilicity, molecular weight, molecular shape, etc., and will be apparent to those having skill in the art. To facilitate purification, it is contemplated that in some embodiments the engineered ketoreductase polypeptides of the present disclosure can be expressed as fusion proteins with purification tags, such as His-tags having affinity for metals, or antibody tags for binding to antibodies, e.g., myc epitope tag.

4.5 Methods of Use

[0167] The engineered ketoreductase polypeptides described herein can be used in processes comprising the conversion of compound (2) to compound (1) as shown in **Scheme 1**, for example in process for manufacturing compound (1), which is used as the active pharmaceutical ingredient, Ezetimibe. Furthermore, the biocatalytic abilities of the non-naturally occurring ketoreductase polypeptides disclosure are not limited to the conversion of compound (2) to compound (1). Additionally, the engineered ketoreductase polypeptides described herein can be used for the conversion of analogs of compound (2) to the corresponding chiral alcohol analogs of compound (1) in diastereomeric excess.

[0168] In some embodiments, the disclosure provides methods for preparing compound (1) or an analog of compound (1) in diastereomeric excess comprising: contacting compound (2) or an analog of compound (2) with an engineered polypeptide of the present disclosure (e.g., as described in Table 2 and elsewhere herein) in the presence of NADPH or NADH cofactor under suitable reaction conditions. Suitable reactions conditions for the conversion of compound (2) to compound (1), or the conversion of an analog of compound (2) to the corresponding analog of compound (1), using the engineered polypeptides of the present disclosure are described in greater detail below and some exemplary suitable reaction conditions also are provided in the Examples.

[0169] The engineered polypeptides of the present disclosure of improved enzymatic properties for the conversion of compound (2) to compound (1) relative to the naturally occurring ketoreductase polypeptide of SEQ ID NO: 2, including increased conversion rates, increased stereoselectivity (resulting in compound (1) in greater diastereomeric excesses), increased solvent stability, and increased thermal stability. Accordingly, it is contemplated that any of the engineered polypeptides disclosed herein may be used in improved methods that comprise the conversion of compound (2) to compound (1). For example, in some embodiments, the methods of the present disclosure can be carried out wherein the engineered polypeptide is selected from an amino acid sequence having at least about 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or more identity to SEQ ID NO: 2, which further comprises the combination of residue differences compared to SEQ ID NO: 2 of any one

of engineered polypeptides disclosed in Table 2 (e.g., even-numbered SEQ ID NOs: 4-168). In some embodiments, the any one or more of the polypeptides of SEQ ID NO: 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, or 168 may be used in the methods disclosed herein.

[0170] The present disclosure also contemplates ranges of suitable reaction conditions that can be used in the methods, including but not limited to ranges of pH, temperature, buffer, solvent system, substrate loading, polypeptide loading, cofactor loading, atmosphere, and reaction time. The present disclosure also contemplates that the methods comprising the biocatalytic conversion of compound (2) to compound (1) using an engineered polypeptide of the disclosure can further comprise chemical steps of compound (1) product work-up, extraction, isolation, purification, and/or crystallization, each of which can be carried out under a range of conditions.

[0171] In some embodiments, the methods for preparing compound (1) of the present disclosure can be carried out wherein the reaction conditions comprise compound (2) substrate loading of at least about 20 g/L, about 40 g/L, about 50 g/L, about 75 g/L, about 100 g/L, about 200 g/L, about 250 g/L, about 300 g/L, about 400 g/L, or even greater. In certain embodiments, methods for preparing compound (1) of the present disclosure can be carried out wherein the reaction conditions comprise compound (2) substrate loading of about 50-100 g/L, about 50-200 g/L, about 50-300 g/L, about 50-400 g/L, about 100 g/L, about 200 g/L, about 300 g/L or about 400 g/L. The values for substrate loadings provided herein are based on the molecular weight of compound (2), however it also contemplated that the equivalent molar amounts of various hydrates and salts of compound (2) also can be used in the methods.

[0172] The improved enzymatic activity of the engineered polypeptides of the present disclosure in the conversion of compound (2) to compound (1) provides for methods wherein higher percentage conversion can be achieved with lower concentrations of the engineered polypeptide. The use of lower concentration of the engineered polypeptide in a method comprising a conversion of compound (2) to compound (1) also reduces the amount of residual protein that may need to be removed in subsequent steps for purification of compound (1). In some embodiments, the methods for preparing compound (1) of the present disclosure can be carried out wherein the reaction conditions comprise an engineered polypeptide concentration of about 0.1-3.0 g/L, about 0.5-2.75 g/L, about 1.0-2.5 g/L, about 1.5-2.5 g/L, about 3 g/L, about 2 g/L, about 1.5 g/L, about 1.0 g/L, about 0.75 g/L, or even lower concentration.

[0173] In certain embodiments, the temperature of the suitable reaction conditions can be chosen to maximize the reaction rate at higher temperatures while maintaining the activity of the enzyme for sufficient duration for efficient conversion of the substrate to the product. Where higher temperatures are used, polypeptides with increased thermostability can be selected to carry out the process.

[0174] The engineered polypeptides of the present disclosure have increased thermal stability relative to the naturally occurring ketoreductase polypeptide of SEQ ID NO: 2. This allows the engineered polypeptides to be used in methods for converting compound (2) to compound (1) at higher temperatures which can result in increased conversion rates and improved substrate solubility characteristics for the reaction, although substrate or product degradation at higher temperatures can contribute to decreased process yields. In certain embodiments, the method can be carried out wherein the reaction conditions comprise a temperature of about 20°C to about 40°C, about 23°C to about 37°C, about 25°C to about 35°C, about 26°C to about 32°C, or about 28°C to about 30°C. In certain embodiments, the temperature during the enzymatic reaction can be maintained at ambient (e.g., 25°C), 27°C, 30°C, 32°C, 35°C, 37°C, 40°C; or in some embodiments adjusted over a temperature profile during the course of the reaction.

[0175] In certain embodiments, the methods for preparing compound (1) of the present disclosure the pH of the reaction mixture may be maintained at a desired pH or within a desired pH range by the addition of an acid or a base during the course of the reaction. In certain embodiments, the pH of the reaction mixture may change or be changed during the course of the reaction. Thus, it is contemplated that in some embodiments the pH may be controlled by using an aqueous solvent that comprises a buffer. Suitable buffers to maintain desired pH ranges are known in the art and include, for example, phosphate buffer, triethanolamine buffer, and the like. Combinations of buffering and acid or base addition may also be used.

[0176] In certain embodiments, the methods for preparing compound (1) of the present disclosure can be carried out wherein the reaction conditions comprise a pH of about 6.0 to about 7.5, a pH of about 6.25 to about 7.25, a pH of about 6.5 to about 7.25, a pH of about 6.6 to about 7.25, a pH of about 6.6 to about 7.0, a pH of about 6.75 to about 7.25, or a pH of about 6.75. Below pH 6.5 the rate of the biocatalytic conversion of compound (2) to compound (1) slows down and consequently a longer overall reaction time (e.g., > 24 h) may be needed to achieve a high level of conversion (e.g., > 97%). Also, NADP⁺ cofactor is less stable below pH 6.25. Above pH 7.25 the degradation of the compound (2) and compound (1) increases, which may result in decreased overall yield and purity of compound (1) made according to the method.

[0177] The methods for preparing compound (1) of the present disclosure are generally carried out in a solvent. Suitable solvents include water, aqueous buffer solutions, organic solvents, and/or co-solvent systems, which generally comprise aqueous solvents and organic solvents.

[0178] In certain embodiments, the methods for preparing compound (1) of the present disclosure can be carried out wherein the reaction conditions comprise a solution comprising an aqueous buffer solution, an organic solvent, or a co-solvent system. In some embodiments, the buffer solution is selected from TEA (e.g., about 0.025 M to about 0.25 M TEA) and potassium phosphate (e.g., about 0.025 M to about 0.25 M phosphate). In certain embodiments, the co-solvent system comprises about 30% (v/v) to about 70% (v/v) of an aqueous buffer solution (e.g., about 0.1 M TEA) and about 70% to about 30% of an organic solvent solution (e.g., IPA and/or toluene). In some embodiments, the reaction conditions comprise water as a suitable solvent with no buffer present.

[0179] The engineered polypeptides of the present disclosure have increased stability to organic solvent relative to the naturally occurring ketoreductase polypeptide of SEQ ID NO: 2. This allows the engineered polypeptides to be used in methods for converting compound (2) to compound (1) in co-solvent systems with higher concentrations of organic solvent which can result in improved product solubility characteristics and increased percent conversion (e.g., 97% or greater conversion of compound (2), at 100 g/L concentration, to compound (1) in 24 h).

[0180] In another embodiment, the co-solvent system comprises an aqueous buffer solution and IPA, wherein the IPA concentration is about 25-75% (v/v), about 35-75% (v/v), about 45-75% (v/v), about 55-75% (v/v), about 60-70% (v/v), about 62-68% (v/v), at least about 25% (v/v), at least about 35% (v/v), at least about 45% (v/v), at least about 55% (v/v), at least about 65% (v/v), about 60% (v/v), about 65% (v/v), or about 70% (v/v). In certain embodiments, the reaction conditions comprise a co-solvent system of 0.1 M TEA buffer and about 60% (v/v) to about 70% (v/v) IPA. In certain embodiments, the reaction conditions comprise a co-solvent system of about 35% (v/v) 0.1 M TEA buffer and about 65% (v/v) IPA.

[0181] In some embodiments, the co-solvent system comprises an aqueous buffer solution, IPA, and another organic solvent, such as toluene. In some embodiments, the co-solvent system comprises about 45-55% (v/v) of an aqueous buffer solution, about 25-35% (v/v) IPA, and about 25-35% (v/v) toluene. In certain embodiments, the co-solvent system comprises about 50% (v/v) of an aqueous buffer solution (e.g., 0.1 M TEA), about 20% (v/v) IPA, and about 30% (v/v) toluene.

[0182] In certain embodiments, the methods comprising the conversion of compound (2) to compound (1) can be carried out wherein the reaction conditions comprise an inert atmosphere (e.g., N₂, Ar, etc.).

[0183] In some embodiments, the methods for preparing compound (1) of the present disclosure can be carried out using a combination of any of the mixture and reaction conditions disclosed above or elsewhere herein, e.g., in the Examples. Accordingly, in some embodiments, the methods of the present disclosure can be carried out wherein the reaction conditions comprise: (1) substrate loading of about 50-200 g/L compound (2); (2) engineered polypeptide concentration of about 1.5-2.5 g/L; (3) NADPH cofactor concentration of about 0.1-0.2 g/L; (4) a co-solvent solution of an aqueous buffer and about 60-70% (v/v) IPA; (5) about pH 6.25-7.5; and (6) temperature of about 25-35°C.

[0184] In some embodiments, the methods for preparing compound (1) of the present disclosure can be carried out wherein the reaction conditions comprise: (1) substrate loading of about 100 g/L compound (2); (2) engineered polypeptide concentration of about 2.0 g/L; (3) NADPH cofactor concentration of about 0.1 g/L; (4) a co-solvent solution of an aqueous buffer of 0.1 M TEA and about 65% (v/v) IPA; (5) about pH 6.75; and (6) temperature of about 30 °C.

[0185] In some embodiments, the methods for preparing compound (1) of the present disclosure can be carried out wherein the reaction conditions comprise: (1) substrate loading of about 50-150 g/L of compound (2); (2) engineered polypeptide concentration of about 2.5-3.5 g/L; (3) NADPH cofactor concentration of about 0.1-0.2 g/L; (4) a co-solvent solution of an aqueous buffer and about 15-25% (v/v) IPA and 25-35% (v/v) toluene; (5) about a pH of 6.6-7.0; and (6) a temperature of about 28-30 °C.

[0186] In some embodiments, the methods for preparing compound (1) of the present disclosure can be carried out wherein the reaction conditions comprise: (1) substrate loading of about 100 g/L of compound (2); (2) engineered polypeptide concentration of about 3.0 g/L; (3) NADPH cofactor concentration of about 0.1 g/L; (4) a co-solvent solution of an aqueous buffer of 0.1 M TEA and about 20% (v/v) IPA and 30% (v/v) toluene; (5) about a pH of 6.6-7.0; and (6) a temperature of about 28-30 °C.

[0187] Generally, in the methods disclosed herein, the biocatalytic reaction with a polypeptide under suitable reaction conditions is allowed to proceed until essentially complete, or near complete, conversion of compound (2) to compound (1) is obtained. This conversion of substrate to product can be monitored using known methods by detecting substrate and/or product. Suitable methods include gas chromatography, HPLC, and the like, and are described in the Examples.

[0188] In some embodiments, the methods for preparing compound (1) of the present disclosure result in at least about 90% conversion of compound (2) at 100 g/L loading to compound (1) in 24 h, when carried out under reaction conditions of: engineered polypeptide concentration of about 1.0-3.0 g/L; NADPH cofactor concentration of about 0.1 g/L; a co-solvent system of at least 65% (v/v) IPA; and a temperature of 30°C. In some embodiments, the methods of the present disclosure when carried out under these reaction conditions (e.g., 100 g/L compound (2) loading) result in at least about 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or greater conversion of compound (2) to compound (1) in 24 h.

[0189] In some embodiments, the methods for preparing compound (1) of the present disclosure when carried out

with 100 g/L compound (2) loading result in an diastereomeric excess of compound (1) of at least 97%, 98, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, 99.6%, 99.7%, 99.8%, or 99.9% in 24 h.

[0190] In carrying out the conversion of compound (2) to compound (1) using the engineered polypeptides in the methods of the present disclosure, it is necessary for an electron donor to be present. Generally, a cofactor is used as the electron donor in the reduction reaction. The cofactor operates in combination with the polypeptides of the disclosure in the process. Suitable cofactors include, but are not limited to, NADP⁺ (nicotinamide adenine dinucleotide phosphate), NADPH (the reduced form of NADP⁺), NAD⁺ (nicotinamide adenine dinucleotide) and NADH (the reduced form of NAD⁺). Generally, the reduced form of the cofactor is added to the reaction mixture. Accordingly, in certain embodiments, the methods of the present disclosure are carried out wherein an electron donor is present selected from NADPH cofactor or NADH cofactor. In certain embodiments, the method can be carried out wherein the reaction conditions comprise an NADH or NADPH cofactor concentration of about 0.03-0.5 g/L, about 0.05-0.3 g/L, about 0.1-0.2 g/L, about 0.5 g/L, about 0.1 g/L, or about 0.2 g/L.

[0191] The reduced NAD(P)H form can be optionally regenerated from the oxidized NAD(P)⁺ form using a cofactor regeneration system. In some embodiments of the process, a cofactor recycling system is used to regenerate cofactor NADPH/NADH form NADP⁺/NAD⁺ produced in the reaction.

[0192] In some embodiments of the process, an optional cofactor recycling system can be used to regenerate cofactor NADPH/NADH form NADP⁺/NAD⁺ produced in the reaction. A cofactor regeneration system refers to a set of reactants that participate in a reaction that reduces the oxidized form of the cofactor (e.g., NADP⁺ to NADPH). Cofactors oxidized by the polypeptide reduction of the keto substrate are regenerated in reduced form by the cofactor regeneration system. Cofactor regeneration systems comprise a stoichiometric reductant that is a source of reducing hydrogen equivalents and is capable of reducing the oxidized form of the cofactor. The cofactor regeneration system may further comprise a catalyst, for example an enzyme catalyst, that catalyzes the reduction of the oxidized form of the cofactor by the reductant. Cofactor regeneration systems to regenerate NADH or NADPH from NAD⁺ or NADP⁺, respectively, are known in the art and may be used in the methods described herein.

[0193] Suitable exemplary cofactor regeneration systems that may be employed include, but are not limited to, glucose and glucose dehydrogenase, formate and formate dehydrogenase, glucose-6-phosphate and glucose-6-phosphate dehydrogenase, a secondary (e.g., isopropanol) alcohol and secondary alcohol dehydrogenase, phosphite and phosphite dehydrogenase, molecular hydrogen and hydrogenase, and the like. These systems may be used in combination with either NADP⁺/NADPH or NAD⁺/NADH as the cofactor. Electrochemical regeneration using hydrogenase may also be used as a cofactor regeneration system. See, e.g., U.S. Pat. Nos. 5,538,867 and 6,495,023. Chemical cofactor regeneration systems comprising a metal catalyst and a reducing agent (for example, molecular hydrogen or formate) are also suitable. See, e.g., PCT publication WO 2000/053731.

[0194] In some embodiments, the cofactor recycling system can comprise glucose dehydrogenase (GDH), which is a NAD⁺ or NADP⁺-dependent enzyme that catalyzes the conversion of D-glucose and NAD⁺ or NADP⁺ to gluconic acid and NADH or NADPH, respectively. Glucose dehydrogenases suitable for use in the practice of the processes described herein include both naturally occurring glucose dehydrogenases, as well as non-naturally occurring glucose dehydrogenases. Naturally occurring glucose dehydrogenase encoding genes have been reported in the literature, e.g., the *Bacillus subtilis* 61297 GDH gene, *B. cereus* ATCC 14579 and *B. megaterium*. Non-naturally occurring glucose dehydrogenases generated using, for example, mutagenesis, directed evolution, and the like and are provided in PCT publication WO 2005/018579, and US publication Nos. 2005/0095619 and 2005/0153417.

[0195] In some embodiments, the co-factor regenerating system can comprise a formate dehydrogenase, which is a NAD⁺ or NADP⁺-dependent enzyme that catalyzes the conversion of formate and NAD⁺ or NADP⁺ to carbon dioxide and NADH or NADPH, respectively. Formate dehydrogenases that are suitable for use as cofactor regenerating systems in the ketoreductase reactions described herein include naturally occurring and non-naturally occurring formate dehydrogenases. Suitable formate dehydrogenases are described in PCT publication WO 2005/018579. Formate may be provided in the form of a salt, typically an alkali or ammonium salt (for example, HCO₂Na, KHCO₂NH₄, and the like), in the form of formic acid, typically aqueous formic acid, or mixtures thereof. A base or buffer may be used to provide the desired pH.

[0196] In some embodiments, the co-factor regenerating system can comprise a secondary alcohol dehydrogenase, which is an NAD⁺ or NADP⁺-dependent enzyme that catalyzes the conversion of a secondary alcohol and NAD⁺ or NADP⁺ to a ketone and NADH or NADPH, respectively. Secondary alcohol dehydrogenases suitable for use as cofactor regenerating systems in the processes described herein include naturally occurring and non-naturally occurring ketoreductases. Naturally occurring secondary alcohol dehydrogenases include known alcohol dehydrogenases from, *Thermoanaerobium brockii*, *Rhodococcus erythropolis*, *Lactobacillus kefir*, and *Lactobacillus brevis*, and non-naturally occurring secondary alcohol dehydrogenases include engineered alcohol dehydrogenases derived therefrom. In some embodiments, non-naturally occurring ketoreductases engineered for thermo- and solvent stability can be used. Such ketoreductases are described in the present application and the patent publications US 20080318295A1; US 20090093031A1; US 20090155863A1; US 20090162909A1; US 20090191605A1; US 20100055751A1;

WO/2010/025238A2; WO/2010/025287A2; and US 20100062499A1.

[0197] The engineered ketoreductase polypeptides of the present disclosure have improved enzymatic activity for the conversion of IPA to acetone relative to the naturally occurring ketoreductase polypeptide of SEQ ID NO: 2. Accordingly, in carrying out the conversion of compound (2) to compound (1) using the engineered polypeptides in the methods of the present disclosure, the NADPH or NADH cofactor present can be recycled by the engineered polypeptide using IPA as reductant.

[0198] In certain embodiments, the methods comprising the conversion of compound (2) to compound (1) disclosed herein can be carried out without adding NADPH or NADH cofactor during the reaction and without any other enzyme systems present (e.g., glucose dehydrogenase, or formate dehydrogenase).

[0199] In certain embodiments, the methods comprising the use of an engineered polypeptide of the present disclosure for the conversion of compound (2) to compound (1) can be carried out wherein no cofactor recycling enzyme is present other than the engineered polypeptide. For example, the methods comprising of the present disclosure can be carried out wherein the reaction conditions comprise an IPA concentration is about 55-75% (v/v), an NADPH or NADH cofactor loading of about 0.03-0.5 g/L, and wherein no cofactor recycling enzyme is present other than the engineered polypeptide.

[0200] Suitable secondary alcohols useful in cofactor regenerating systems include lower secondary alkanols and aryl-alkyl carbinols. Examples of lower secondary alcohols include isopropanol, 2-butanol, 3-methyl-2-butanol, 2-pentanol, 3-pentanol, 3,3-dimethyl-2-butanol, and the like. In one embodiment, the secondary alcohol is isopropanol. Suitable aryl-alkyl carbinols include unsubstituted and substituted 1-arylethanol.

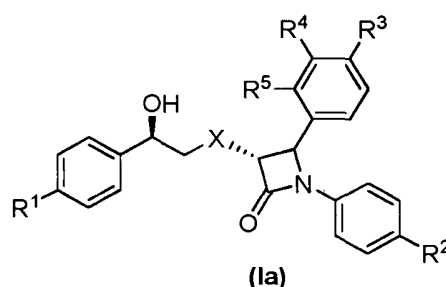
[0201] In some embodiments where the cofactor recycling system produces a volatile product, such as acetone from isopropanol, the volatile product can be removed by sparging the reaction solution with a non-reactive gas or by applying a vacuum to lower the reaction pressure and removing the volatile present in the gas phase. A non-reactive gas is any gas that does not react with the reaction components. Various non-reactive gases include nitrogen and noble gases (e.g., inert gases). In some embodiments, the non-reactive gas is nitrogen gas. For example, acetone formed by oxidation of isopropanol can be removed by sparging with nitrogen gas or applying a vacuum to the reaction solution and removing the acetone from the gas phase by an acetone trap, such as a condenser or other cold trap.

[0202] In the embodiments herein, the polypeptides carrying out the conversion of compound (2) to compound (1) and any additional enzymes of the optional cofactor regeneration system, may be added to the reaction mixture in the form of the purified enzymes, whole cells transformed with gene(s) encoding the enzymes, and/or cell extracts and/or lysates of such cells. The gene(s) encoding the polypeptides disclosed herein and the optional cofactor regeneration enzymes can be transformed into host cells separately or together into the same host cell. Whole cells transformed with gene(s) encoding the engineered ketoreductase enzyme and/or the optional cofactor regeneration enzymes, or cell extracts and/or lysates thereof, may be employed in a variety of different forms, including solid (e.g., lyophilized, spray-dried, and the like) or semisolid (e.g., a crude paste).

[0203] Generally, the order of addition of reactants (e.g., substrate, cofactor, polypeptide) is not critical to the methods of the present disclosure. The reactants may be added together at the same time to a solvent (e.g., monophasic solvent, biphasic aqueous co-solvent system, and the like), or alternatively, some of the reactants may be added separately, and some together at different time points.

[0204] In some embodiments any of the above describe methods for the conversion of compound (2) to compound (1) can be carried out wherein the method comprises contacting an analog of compound (2) with an engineered polypeptide of the present disclosure (e.g., as described in Table 2 and elsewhere herein) in the presence of NADPH or NADH cofactor under suitable reaction conditions, thereby resulting in the preparation of the chiral alcohol of the corresponding analog of compound (1) in diastereomeric excess. Suitable reactions conditions for the conversion of analogs of compound (2) to the chiral alcohol of the corresponding analogs of compound (1) can be the same as used for compound (2) or determined by the ordinary artisan based on the known properties of the analog compounds and routine experimentation.

[0205] In some embodiments, the analogs of compound (1) prepared using the above described methods include the analogs of compound (1) comprising the compound of Formula Ia shown below.



wherein,

X is selected from C or S;

R¹ is selected from -H, -F, -Cl, -Br, or -I;

R² is selected from -H, -F, -Cl, -Br, -I, -CN, -OH (optionally protected with a hydroxyl protecting group), -CH₂NH₂ (optionally protected with a nitrogen protecting group), and any one of the following optionally substituted groups: alkyl, alkoxy, alkenyl, alkenoxy, alkynyl, alkynoxy, cycloalkyl, aryl, heteroaryl, or heterocycle;

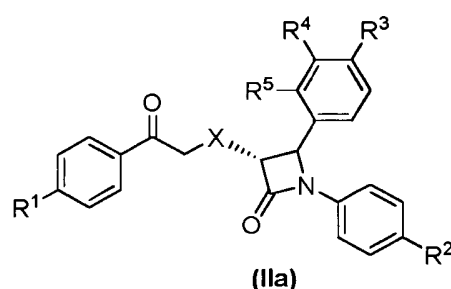
R³ is selected from -H, -F, -Cl, -Br, -I, -CN, -OH (optionally protected with a hydroxyl protecting group), -CH₂NH₂ (optionally protected with a nitrogen protecting group), and any one of the following optionally substituted groups: alkyl, alkoxy, alkenyl, alkenoxy, alkynyl, alkynoxy, cycloalkyl, aryl, heteroaryl, or heterocycle;

R⁴ is selected from -H, -F, -Cl, -Br, -I, -CN, -OH (optionally protected with a hydroxyl protecting group); and

R⁵ is selected from -H, -F, -Cl, -Br, -I, -CN, -OH (optionally protected with a hydroxyl protecting group), -CH₂NH₂ (optionally protected with a nitrogen protecting group), and any one of the following optionally substituted groups: alkyl, alkoxy, alkenyl, alkenoxy, alkynyl, alkynoxy, cycloalkyl, aryl, heteroaryl, or heterocycle.

[0206] Examples of hydroxyl protecting groups and nitrogen protecting groups that may be the R group of compounds of Formula **IIa** undergoing the biocatalytic methods of the present disclosure can be found in P.G.M. Wuts and T. W. Greene, "Greene's Protective Groups in Organic Synthesis - Fourth Edition," John Wiley and Sons, New York, N.Y., 2007, Chapter 7 ("Greene").

[0207] Accordingly, in some embodiments the present disclosure provides a method of preparing a compound of Formula **Ia** comprising: contacting a compound of Formula **IIa**



wherein,

X, R¹, R², R³, R⁴, and R⁵, are defined as above for Formula **Ia**, with an engineered polypeptide of the present disclosure (e.g., as described in Table 2 and elsewhere herein) in the presence of NADPH or NADH cofactor under suitable reaction conditions.

[0208] In some embodiments, the present disclosure provides a method of preparing a compound of Formula **Ia** in which R² and R³ are each independently selected from -H, -F, -Cl, -Br, -I, -CN, -OH (optionally protected with a hydroxyl protecting group), -CH₂NH₂ (optionally protected with a nitrogen protecting group), and any one of the following optionally substituted groups: -(C₁-C₆)alkyl, -(C₁-C₆)alkoxy, -(C₁-C₆)alkenyl, -(C₁-C₆)alkenoxo, -(C₁-C₆)alkynyl, -(C₁-C₆)alkynoxy, -(C₁-C₆)cycloalkyl, or a heterocycle, having from 1 to 4 carbon atoms and 1 to 2 hetero atoms.

[0209] In some embodiments, the present disclosure provides a method of preparing a compound of Formula **Ia** in which R² and R³ are each independently selected from -H, -F, -Cl, -Br, -I, -CN, -OH (optionally protected with a hydroxyl protecting group), -CH₂NH₂ (optionally protected with a nitrogen protecting group), and any one of the following optionally substituted groups: -(C₁-C₄)alkyl, -(C₁-C₄)alkoxy, -(C₁-C₄)alkenyl, -(C₁-C₄)alkenoxo, -(C₁-C₄)alkynyl, -(C₁-C₄)alkynoxy, -(C₁-C₄)cycloalkyl, or a heterocycle, having from 1 to 3 carbon atoms and 1 to 2 hetero atoms.

[0210] In some embodiments, the present disclosure provides a method of preparing a compound of Formula **Ia** in which X is C, R² and R³ are each independently selected from -H, -F, -Cl, -Br, -I, -CN, -OH (optionally protected with a hydroxyl protecting group), and -CH₂NH₂ (optionally protected with a nitrogen protecting group).

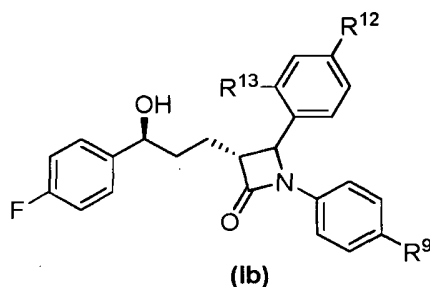
[0211] In some embodiments, the present disclosure provides a method of preparing a compound of Formula **Ia** in which X is C, R² and R³ are each independently selected from the following optionally substituted groups: -(C₁-C₄)alkyl, -(C₁-C₄)alkoxy, -(C₁-C₄)alkenyl, -(C₁-C₄)alkenoxo, -(C₁-C₄)alkynyl, and -(C₁-C₄)alkynoxy. In some embodiments, R² and R³ are optionally substituted with one or more -OH groups.

[0212] In some embodiments, the present disclosure provides a method of preparing a compound of Formula **Ia** in which X is C, any one or more of R², R³, or R⁵ is an -OH or an -OH group protected with a hydroxyl protecting group.

In some embodiments of the methods, the hydroxyl protecting group of R², R³, and/or R⁵ is selected from the group consisting of benzyl, acetyl, benzoyl, *tert*-butyloxycarbonyl, silyl, *tert*-butyldiphenylsilyl, trimethylsilyl, *para*-methoxybenzyl, benzylidene, dimethylacetal, and methoxy methyl. In some embodiments, the silyl group is -Si-(R^a)(R^b)(R^c) and R^a, R^b, and R^c are each independently selected from the group consisting of C₁-C₆ alkyl, phenyl, acetyl, and benzyl groups.

[0213] In some embodiments, the present disclosure provides a method of preparing a compound of Formula **1a** in which R¹ is -F, R² is -F, R⁴ is -H, and/or R⁵ is -H.

[0214] In some embodiments, the analogs of compound (1) prepared using the above described methods include the anti-hypercholesterolemic compounds described in PCT publication WO 2008/085300A1 based on compounds of Formula **1b**,



wherein,

R⁹ is selected from the group consisting of chloro, fluoro, -C≡C-(C₁-C₆)alkyl-NR¹⁰R¹¹, -(CH₂)_xCH=CH-(C₁-C₆)alkyl-NR¹⁰R¹¹, (C₁-C₈)alkyl-NR¹⁰R¹¹, -C≡C-(C₁-C₄)alkyl-CH-(CH₂-NR¹⁰R¹¹)₂, -(C₁-C₆)alkyl-CH-(CH₂-NR¹⁰R¹¹)₂, -C≡C-(C₁-C₆)alkyl-R^{11a}, -(CH₂)_xCH=CH-(C₁-C₆)alkyl-R^{11a}, -(C₁-C₈)alkyl-R^{11a}, -C≡C-(C₁-C₆)alkyl, -(CH₂)_xCH=CH-(C₁-C₆)alkyl, -(C₁-C₈)alkyl, -(C₁-C₁₅)alkynyl mono- or poly-substituted with -OH and optionally substituted with R¹⁴, -(C₁-C₁₅)alkenyl mono- or poly-substituted with -OH and optionally substituted with R¹⁴, -(C₁-C₁₅)alkyl mono- or poly-substituted with -OH and optionally substituted with R¹⁴ and x is an integer selected from 0, 1 and 2;

R¹⁰ is independently selected at each occurrence from the group consisting of -H and -(C₁-C₃)alkyl;

R¹¹ is independently selected at each occurrence from the group consisting of -H, -(C₁-C₃)alkyl, -C(O)-(C₁-C₃)alkyl, -C(O)-NR¹⁰R¹⁰, -SO₂-(C₁-C₃)alkyl and -SO₂-phenyl;

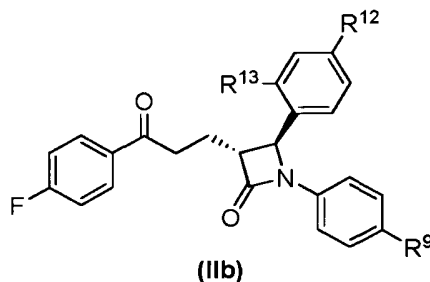
R^{11a} is selected from the group consisting of -C(O)-NR¹⁰R¹⁰, -SO₂-(C₁-C₃)alkyl and -SO₂-phenyl;

R¹² is selected from the group consisting of -(C₂-C₁₅)alkynyl mono- or poly-substituted with -OH and optionally substituted with R¹⁴, -(C₂-C₁₅)alkenyl mono- or poly-substituted with -OH and optionally substituted with R¹⁴, -(C₂-C₁₅)alkyl mono- or poly-substituted with -OH and optionally substituted with R¹⁴;

R¹³ is selected from the group consisting of -H and -OH; and

R¹⁴ is a sugar residue optionally substituted with -COOH, -COO-(C₁-C₃)alkyl and -(C₁-C₃)alkyl-OH; provided that when R⁹ is selected from the group consisting of -C≡C-(CH₂)₁₋₆-NR¹⁰R¹¹, -CH=CH-(CH₂)₁₋₆-NR¹⁰R¹¹ and -(CH₂)₁₋₆-NR¹⁰R¹¹, then R¹² is not selected from the group consisting of -(C₁-C₁₅)alkyl mono- or poly-substituted with -OH, -CH=CH-(C₁-C₃)alkyl mono- or poly-substituted with -OH, -C≡C-(C₁-C₃)alkyl mono- or poly-substituted with -OH, and -(CH₂)₀₋₁-C(=CH₂)-CH₂OH.

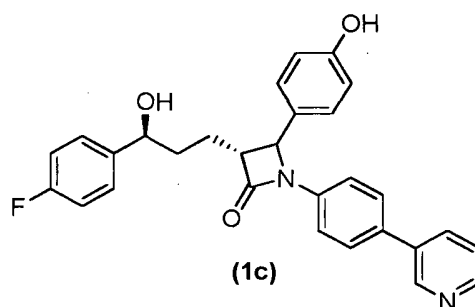
[0215] Accordingly, in some embodiments the present disclosure provides a method of preparing a compound of Formula **1b** comprising: contacting a compound of Formula **11b**,



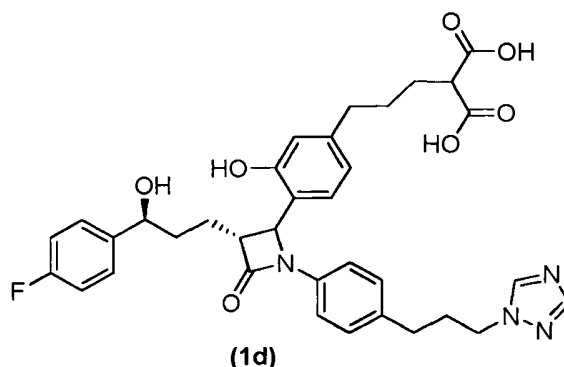
wherein,

R⁹, R¹⁰, R¹¹, R^{11a}, R¹², R¹³, and R¹⁴ are defined as above for Formula **1b**, with an engineered polypeptide of the present disclosure (e.g., as described in Table 2 and elsewhere herein) in the presence of NADPH or NADH cofactor under suitable reaction conditions.

[0216] In some embodiments, the analogs of compound **(1)** prepared using the above described methods include the anti-hypercholesterolemic compounds described in published Japan patent application 2010-83880, including the compound **(1c)** and other related analog compounds having substituted pyridines or other heteroaryl at R⁹ of compound of Formula **(1b)**:

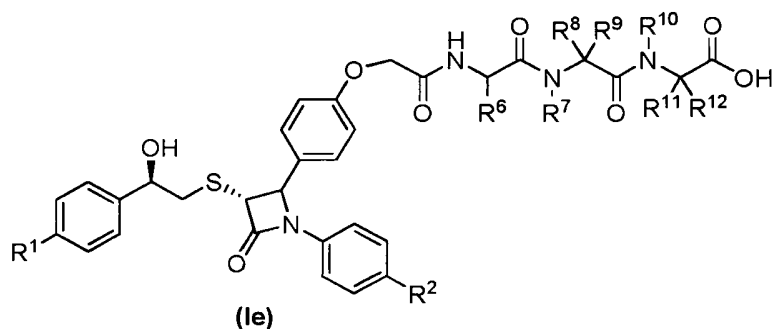


[0217] In some embodiments, the analogs of compound **(1)** prepared using the above described methods include the anti-hypercholesterolemic compounds described in PCT publication WO2010/056788, including the compound **(1d)** and other related analog compounds having substituted alkyl chains at R⁹ and R¹² positions of the compound of Formula **(1b)**:



[0218] In some embodiments, the analogs of compound **(1)** prepared using the above described methods include the anti-hypercholesterolemic compounds described in US published patent application US2010/160282A1 and PCT Publication WO2010/100255.

[0219] In some embodiments, the analogs of compound **(1)** prepared using the above described methods include the anti-hypercholesterolemic compounds described in US published patent application US2010/152156A1, including the compound of Formula **(1e)** and other related analog compounds having a sulfur atom at position X of the compound of Formula **(1a)**:



wherein:

R¹ is hydrogen, alkyl, halo, C₁₋₆alkoxy or C₁₋₆ alkylS-;

R² is hydrogen, C₁₋₆ alkyl, halo or C₁₋₆alkoxy;

R⁶ is hydrogen, C₁₋₆ alkyl, C₃₋₆ cycloalkyl or aryl;

R⁸, R⁹, R¹¹ and R¹² are independently hydrogen, a branched or unbranched C₁₋₆ alkyl, C₃₋₆ cycloalkyl or aryl; wherein said C₁₋₆ alkyl may be optionally substituted by one or more hydroxy, amino, guanidino, cyano, carbamoyl, carboxy, C₁₋₆alkoxy, aryl C₁₋₆alkoxy, (C₁₋₄alkyl)₃Si, N-(C₁₋₆alkyl)amino, N,N-(C₁₋₆alkyl)₂-amino, C₁₋₆alkylS(O)_a, C₃₋₆cycloalkyl, aryl or aryl C₁₋₆ alkyl-S(O)_a, wherein a is 0-2; and wherein any aryl group may be optionally substituted

by one or two substituents selected from halo, hydroxy, C₁₋₆alkyl, C₁₋₆alkoxy, or cyano;

R⁷ and R¹⁰ is hydrogen, C₁₋₆ alkyl, or arylC₁₋₆ alkyl;

and, wherein R⁸ and R⁹ may form a ring with 2-7 carbon atoms and wherein R⁷ and R⁸ may form a ring with 3-6 carbon atoms; or a pharmaceutically acceptable salt, solvate, solvate of such a salt or a prodrug thereof.

5. EXAMPLES

[0220] Various features and embodiments of the disclosure are illustrated in the following representative examples, which are intended to be illustrative, and not limiting.

Example 1: Wild-type ketoreductase gene acquisition and construction of expression vectors.

[0221] The wild-type ketoreductase gene from *L. kefir* (SEQ ID NO: 1) was designed for expression in *E. coli* using standard codon optimization. (Codon-optimization software is reviewed in e.g., "OPTIMIZER: a web server for optimizing the codon usage of DNA sequences," Puigbò et al., Nucleic Acids Res. 2007 Jul; 35(Web Server issue): W126-31. Epub 2007 Apr 16.) Genes were synthesized using oligonucleotides composed of 42 nucleotides and cloned into expression vector pCK110900 (vector depicted as FIG. 3 in US Patent Application Publication 20060195947) under the control of a *lac* promoter. The expression vector also contained the P15a origin of replication and the chloramphenicol resistance gene. Resulting plasmids were transformed into *E. coli* W3110 (fhu-) using standard methods. Polynucleotides encoding the engineered ketoreductase polypeptides were also cloned into vector pCK110900 for expression in *E. coli* W3110.

[0222] The engineered ketoreductase polypeptide of SEQ ID NO: 2 which was derived previously based on directed evolution of a codon-optimized gene encoding the wild-type ketoreductase of *Lactobacillus kefir* (Genbank acc. No. AAP94029.1; GI: 33112056). SEQ ID NO: 2 has 19 amino acid residue differences relative to the WT ketoreductase (D3N, G7S, L17Q, V95L, S96Q, G117S, Q127R, E145S, F147L, T152M, L153V, L176V, Y190C, D198K, L199D, E200P, K211R, I223V, and A241 S). The polypeptide of SEQ ID NO: 2 was found to be able to convert compound (2) to compound (1) in > 99% ee and with greater than 50% conversion rate in 20 h while converting IPA to acetone to recycle the NADP⁺ co-factor in 20% IPA (i.e., without a secondary enzyme for cofactor recycling) under initial screening conditions (4 g/L compound (2) substrate; 0.5 g/L NADP, 100 mM TEA, pH 7.0, 1 mM MgSO₄, 25°C). The polypeptide SEQ ID NO: 2 was used as the starting backbone for subsequent rounds of evolution. Multiple rounds of directed evolution of the gene encoding SEQ ID NO: 2 (i.e., SEQ ID NO: 1) were carried out. Each round used the gene encoding the most improved engineered polypeptide from each round as the parent "backbone" sequence for the subsequent round of evolution. The resulting engineered ketoreductase polypeptide sequences and specific mutations and relative activities are listed in Table 2.

Example 2: Production of engineered ketoreductase polypeptides

[0223] The engineered ketoreductase polypeptides of the disclosure were produced in *E. coli* W3110 as an intracellular protein expressed under the control of the *lac* promoter. The polypeptide accumulates primarily as a soluble cytosolic active enzyme. A shake-flask procedure is used to generate engineered polypeptide powders that can be used in activity assays or biocatalytic process disclosed herein.

[0224] Fermentation for shake flask powders: A single microbial colony of *E. coli* containing a plasmid encoding an engineered ketoreductase of interest is inoculated into 50 mL Luria Bertani broth containing 30 µg/ml chloramphenicol and 1% glucose. Cells are grown overnight (at least 16 hours) in an incubator at 30°C with shaking at 250 rpm. The culture is diluted into 250 mL Terrific Broth (12 g/L bacto-tryptone, 24 g/L yeast extract, 4 mL/L glycerol, 65 mM potassium phosphate, pH 7.0, 1 mM MgSO₄) containing 30 µg/ml chloramphenicol, in a 1 liter flask to an optical density at 600 nm (OD₆₀₀) of 0.2 and allowed to grow at 30°C. Expression of the ketoreductase gene is induced by addition of isopropyl-β-D-thiogalactoside ("IPTG") to a final concentration of 1 mM when the OD₆₀₀ of the culture is 0.6 to 0.8 and incubation is then continued overnight (at least 16 hours). Cells are harvested by centrifugation (5000 rpm, 15 min, 4°C) and the supernatant discarded.

[0225] Production of ketoreductase shake-flask powders: The cell pellet is resuspended with an equal volume of cold (4°C) 100 mM triethanolamine (chloride) buffer, pH 7.0 (optionally including 2 mM MgSO₄), and harvested by centrifugation as above. The washed cells are resuspended in two volumes of the cold triethanolamine (chloride) buffer and passed through a French Press twice at 12,000 psi while maintained at 4°C. Cell debris is removed by centrifugation (9000 rpm, 45 minutes, 4°C). The clear lysate supernatant was collected and stored at -20°C. Lyophilization of frozen clear lysate provides a dry shake-flask powder of crude ketoreductase polypeptide. Alternatively, the cell pellet (before or after washing) can be stored at 4°C or -80°C.

[0226] Fermentation for production downstream process (DSP) powders: Larger-scale (~100 - 120 g) fermentation of the engineered ketoreductases for production of DSP powders can be carried out as a short batch followed by a fed batch process according to standard bioprocess methods. Briefly, ketoreductase expression is induced by addition of IPTG to a final concentration of 1 mM. Following fermentation, the cells are harvested and resuspended in 100 mM Triethanolamine-H₂SO₄ buffer, then mechanically disrupted by homogenization. The cell debris and nucleic acid are flocculated with polyethylenimine (PEI) and the suspension clarified by centrifugation. The resulting clear supernatant is concentrated using a tangential cross-flow ultrafiltration membrane to remove salts and water. The concentrated and partially purified enzyme concentrate can then be dried in a lyophilizer and packaged (e.g., in polyethylene containers).

Example 3: Activity Assay of Engineered Ketoreductase Polypeptides

[0227] High-throughput growth & expression: Picked and grown using standard KRED protocol for W3110 with direct induction: (1) Master growth = single colonies picked from agar Q-trays by Q-bot and grown overnight in LB media containing 1% glucose and 30 µg/mL CAM, 30 °C, 200 rpm, 85% humidity. (2) Subculture = 20 µL of overnight growth transferred to a deep well plate containing 380 µL 2xYT growth media containing 30 µg/mL CAM, 1 mM IPTG, 1 mM MgSO₄, and incubated for ~18 h at 30 °C, 200 rpm, 85% humidity. (3) Cell culture centrifuged at 4000 rpm, 4 °C for 10 min., and used media discarded. Cell pellets resuspended in between 200 - 400 µL lysis buffer (100 mM TEA buffer, pH7.0, containing 1 mM MgSO₄, 400 µg/mL PMBS and 500 µg/mL Lysozyme).

[0228] High Throughput Screening Procedure: 60 - 140 µL of a 5.7 - 100 g/L solution of the substrate in either a mixture of toluene:IPA:acetone (v/v/v ratio of 5:9:1 - 15:9:1) or IPA:acetone (v/v ratio of 49:1) was added to each well of a Costar™ deep-well 96-well plate. Subsequently, 40 - 120 µL of a 0.25 - 1.25 g/L solution of NADP in 100 mM TEA buffer, pH 7.0 containing 1 mM MgSO₄ was also added. Finally, 20 µL of a freshly prepared suspension of lysed cells in lysis buffer (either concentrated or diluted up to 20-fold in 100 mM TEA buffer pH 7.0 containing 1 mM MgSO₄) was added to make the total volume in each well 200 µL ([substrate] = 4 - 80 g/L, [NADP] = 0.1 - 0.5 g/L, Solvent = either toluene:IPA:acetone:buffer (relative % volumes of 10:18:2:70 or 30:18:2:50) or IPA:acetone:buffer (relative % volumes of 65:35 or 70:30)). The plate was then heat sealed and shaken for 2 or 24 h at RT, 30°C, or 37 °C before 0.8 mL of acetonitrile was added to each well to quench the reaction.

[0229] The specific conditions of the High Throughput Screening procedure can be varied in order to identify variant polypeptides having amino acid differences providing different improved properties relative to the selected reference polypeptide. Typically, the stringency of screening conditions are increased through the course of the directed evolution of the variant polypeptides. Conditions that can be varied include substrate concentration, cofactor concentration, solvent conditions, temperature, and total reaction time. Exemplary modifications of the screening conditions are noted in Table 2.

[0230] Analytical method used for activity assay: The plate containing the reactions quenched with acetonitrile were heat sealed, and shaken for 5 minutes, prior to being centrifuged at 4,000 rpm for 10 min. 200 µL of the supernatant was then transferred to a Costar™ round bottom 96-well plate and heat sealed prior to HPLC analysis. HPLC was performed using a C-18 Symmetry 100 x 4.6 mm, 5 µm column, with isocratic elution of a 66% MeCN:34% H₂O solvent mixture at a flow of 2.5 mL/min. Both substrate and product were detected by UV absorbance at 254 nm.

Example 4: Biocatalytic Process I for Preparation of Compound (1) (Ezetimibe) from Substrate Compound (2)

[0231] This example illustrates a first biocatalytic process using an engineered ketoreductase polypeptide of the disclosure to prepare Ezetimibe (compound (1)) on a 10 g scale. The biocatalytic reaction is carried out in an aqueous co-solvent system of TEA buffer (100 mM, pH 7), 30% toluene, 20% IPA, and a substrate loading is 100 g/L. The engineered ketoreductase (polypeptide of SEQ ID NO: 80 at 3 g/L loading) uses the cofactor NADPH (0.1 g/L loading) as a reducing agent, which is oxidized to NADP⁺ during the reaction. The engineered ketoreductase also acts as the secondary alcohol dehydrogenase in an in-situ "recycling system" to regenerate the reduced form of the cofactor through the oxidation of the IPA co-solvent to acetone. The product in the biocatalytic reaction is extracted into THF and solvent swap with toluene provides the desired crude product of compound (1), which is then further crystallized from THF/toluene.

[0232] Preparation of compound (2) substrate: Compound (2) for use as substrate in the biocatalytic reaction can be prepared by oxidation of samples of the Ezetimibe API (compound (1)) according to the following procedure. An oven dried 2-neck 500 mL RB flask equipped with a thermocouple, a magnetic stir bar, and a nitrogen gas inlet was charged

with the white powder of compound **(1)** (32.06 g, 78.3 mmol) and N-methyl morpholine oxide (NMO) (18.3 g, 156.6 mmol). 300 mL of anhydrous dichloromethane was added, affording a clear yellow solution. Oven dried activated 4 Å molecular sieves (35 g) was added, and the solution was cooled to 8 °C (internal temperature) using an ice/NaCl bath. Tetrapropylammonium perruthenate (TPAP) (2.75 g, 7.83 mmol) was added in one portion to the flask. The internal temperature rose to 15.4 °C then slowly dropped to 8 °C. The ice bath was removed, and the reaction mixture was allowed to stir at 25 °C for 2.5 h. The dark brown solution was filtered through a 4" bed of Celite and rinsed with dichloromethane (1.5 L) and diethyl ether (500 mL). The filtrate was monitored by TLC to ensure all ketone product had been eluted. The sample was concentrated under reduced pressure and purified by column chromatography with 25% EtOAc in heptane yielding compound **(2)** as an off-white solid, 21.01 g (68% Th), 100% pure at UV₂₅₄. HPLC analysis of a compound **(2)** sample in acetonitrile (~1-5 mg/mL) can be run on an Eclipse XDB-C18 column under the following conditions: T = 35°C; mobile phase A = water + 0.1% TFA; mobile phase B = acetonitrile + 0.1% TFA; run time = 10 min; 0-4.5 min = 25% to 90% B; 4.5-5.25 min = 90% B; 5.25-6 min = 90%-25% B; post-time = 1 min at 25% B. UV detection at 214 and 254 nm. Compound **(2)** retention time = 5.85 min.

[0233] Biocatalytic reaction procedure: A 250 mL round bottomed flask was equipped with overhead stirrer and internal thermometer. The reactor was charged sequentially with 300 mg DSP powder of engineered ketoreductase polypeptide of SEQ ID NO: 80, 50.0 mL 100 mM TEA buffer (pH 7), 10.0 mg NADP⁺ dissolved in buffer, 10.0 g of compound **(2)** dissolved in 30 mL toluene, and 20 mL IPA. The resulting slurry reaction was heated to 30 °C (internal temperature), stirring at ~500 rpm. The final temperature was reached within 15 min. The reaction was run at a starting pH of 7 and the pH remained constant throughout the reaction time. The reaction course was followed periodically by taking samples out of the reaction mixture, quenching, and analyzing as described in HPLC Method 1. At 24 h, the reaction solution was a white suspension and the Method 1 in-process analysis indicated 82% conversion. At 48 h, in-process analysis indicated 94% conversion, and the reaction was cooled to 25°C.

[0234] Crude product work-up procedure: THF (60 mL) was charged to the reaction mixture at room temperature and agitated at 250 rpm for 15 minutes. Phases were allowed to separate and the aqueous layer removed. The THF phase was collected separately. The purity of the product in the THF phase was determined to 94.6 % according to HPLC (Method 1). Toluene (60 mL) was added, and the resulting solution was concentrated to approximately 60 mL on rotary evaporator at 40°C and incrementally reducing the pressure to 70 Torr. Toluene (60 mL) was added again, and the resulting hazy solution was concentrated to approx 90 mL on rotary evaporator at 40°C and incrementally reducing the pressure to 70 Torr. At this stage product precipitated as white solid and GC analysis of the organic (THF) layer indicated that ≤2.0% THF remained. The precipitated product was recovered by filtration and the residue was washed with 1x15 mL of toluene and dried under vacuum (approx 20 mm Hg) for 24 hours. This provided: 8.75 g (90% yield) chiral alcohol of compound **(1)** as a white solid; chemical purity of 97.2% (AUC, HPLC Method 1).

[0235] Crystallization procedure: To a suspension of 8.0 g crude compound **(1)** in toluene (80 mL) at 82°C (internal), THF (10 mL) was added slowly while stirring. The slurry became clear solution at the end of the addition. The solution was allowed to cool to room temperature (25°C) over a period of 12 hours while stirring magnetically (100 RPM). The resulting white precipitate was filtered under reduced pressure. The white residue was washed once with cold toluene (10 mL) and dried under vacuum (~20 mm Hg) for 24 hours. This provided: 6.3 g (68% yield) of compound **(1)** in a single crop as a white solid; chemical purity = 99.9% (AUC, HPLC Method 1); chiral purity >99.9% d.e. (Method 2).

[0236] Analytical methods using in the process of Example 4: Samples were analyzed for percent conversion and/or diastereomeric purity using HPLC according to Method 1 or Method 2 as described below. HPLC samples were prepared as follows: 10 µL are taken from the reaction suspension via pipette, dissolved in 1 mL of acetonitrile, and injected neat into the HPLC according the Method 1 or Method 2 parameters.

[0237] HPLC Method 1 parameters for monitoring biocatalytic reaction progress are shown in Table 4.

Table 4: Method 1 HPLC parameters

Instrument	Agilent HPLC 1200 series
Column	Symmetry C18 4.6x100mm
Mobile Phase	70% Acetonitrile + 0.1% TFA, 30% Water + 0.1% TFA isocratic
Flow Rate	1.0 ml/min
Detection Wavelength	280.0 nm
Detector Temperature	45°C
Injection Volume	10 µl
Run time	4.0 min
Retention times	Product [Compound (1)]: 2.03 min

EP 2 566 497 B1

(continued)

	Substrate [Compound (2)]: 2.60 min Toluene: 3.67 min
Response factor (Substrate/Product)	1.4

[0238] Method 2 HPLC parameters for determining diastereomeric purity of biocatalytic reaction product are shown in Table 5.

Table 5: Method 2 HPLC parameters

Instrument	Agilent HPLC 1200 series
Column	Chiralpak AD-H 4.6x250mm (5um)
Mobile Phase	80% Heptane / 20% EtOH (0-18min, isocratic) 50% Heptane / 50% EtOH (18.5-33min, isocratic)
Flow Rate	1.0 ml/min
Detection Wavelength	230.0 nm
Detector Temperature	20°C
Injection Volume	10 µl
Run time	45.0 min
Retention times	Substrate [Compound (2)]: 31.56 min Product [Compound (1)]: 16.44 min (R,R,S) diastereomer: 15.05 min

Example 5: Process II for Preparation of Ezetimibe from Compound (2)

[0239] This example illustrates a second biocatalytic process using an engineered ketoreductase polypeptide of the disclosure to prepare Ezetimibe (Compound (1)) on a 20 g scale.

[0240] Biocatalytic reaction procedure: A 500 mL jacketed reactor was charged sequentially with the following: 20 g compound (2) (assayed at 88% w/w purity) as solid, 130 mL IPA, 55 mL TEA buffer, 2.0 mL of TEA buffer solution containing 20 mg NADP⁺, 13.0 mL of TEA buffer solution containing 400 mg of the engineered KRED of SEQ ID NO: 168. The resulting reaction mixture was stirred at 30 °C (internal) at ~250 rpm. The pH of the reaction mixture ranges between 6.30 - 6.40 at 30 °C, with an initial buffer pH of 6.75 at RT. The reaction course was followed periodically by taking samples from the reaction mixture, quenching, and analyzing as described in Method 3. Percent conversion at 4 h, 18 h, and 19 h, was 72.0%, 98.2%, and 98.4%, respectively. After in-process analysis (Method 3) indicated maximum possible conversion (at 98.4% conversion) the reaction mixture was taken for the subsequent workup and isolation procedure.

[0241] Product work-up and isolation: Acetone formed during the reaction was distilled under vacuum (40 torr) at 30 °C. To the thick slurry, water (200 mL) was added and the distillation continued at a slightly elevated temperature (40 torr, 40 °C) until about 25% IPA remained relative to the start of distillation. The slurry with the crude product was drained from the reactor. The reactor was washed with another 200 mL of water and drained into the same container from above. The crude product was collected by filtration through a sintered funnel and washed with 50 mL water. The wet cake was dried for 15 h under vacuum (5 torr) at 25 °C. Upon drying, 16.0 g of crude product was obtained 99.0% chemical purity (AUC, HPLC, 99.8% d.e.). Yield of crude product is 90% with respect to the effective loading of keto phenol substrate (17.6 g). The crude product was further purified by recrystallization as described below.

[0242] Recrystallization: A suspension of crude product (10.0 g) in IPA (30.0 mL) was heated to 60 °C (internal) to allow maximum dissolution of product. The hot solution from above was passed through a celite (5.0 g) bed in a sintered funnel. Upon complete filtration, the celite bed was washed with pre-heated IPA (30.0 mL, ~60 °C). Distillation to dryness of the combined filtrates from above showed >9.0 g of product. The white solid was stirred in 30.0 mL IPA and heated to 60 °C (internal) to obtain a clear solution. Water (40.0 mL) was added drop wise to the above solution at 60 °C and the resultant solution was allowed to cool to 25 °C. The crystallized product was filtered through a sintered funnel and dried under vacuum (5 torr, 25 °C) for 15 h. This provided: 8.8 g of chiral alcohol product of compound (1) in a single crop as a white solid; 99.5% chemical purity (AUC, HPLC, 99.9% d.e.). Essentially the only detectable impurity was the

EP 2 566 497 B1

keto phenol substrate, measuring 0.38% (AUC, HPLC).

[0243] Analytical methods used in the process of Example 5: an HPLC method for determination of % conversion (Method 3); and a chiral HPLC method for determination of diastereomeric purity (Method 4).

[0244] Method 3 sample preparation for HPLC: 0.3 mL of reaction mixture was sampled from the stirred reaction suspension via pipette. The appearance of the sample should be as finely dispersed as the reaction mixture itself. The sample was fully dissolved in 25 mL of methanol or acetonitrile. Injection is neat into the HPLC.

[0245] The HPLC parameters used for determination of percent conversion according to Method 3 are shown in Table 6.

Table 6

Instrument	Varian 920-LC series
Column	Alltima C18, 53 x 7 mm, 3 μ m with guard column (P/N: 50605)
Mobile Phase	60% Acetonitrile, 40% Water (Isocratic)
Flow Rate	1.3 mL/min
Detection Wavelength	254.0 nm
Column Temperature	Ambient
Injection Volume	10 μ L
Run time	5.0 min
Retention times	Product Compound (1) : 2.67 min Keto phenol Substrate Compound (2) : 3.92 min
Response Factor (Substrate/Product)	1.46

[0246] Method 4 HPLC sample preparation of in-process sample: 10 μ L of reaction mixture is sampled from the stirred reaction suspension via micropipette and added to 1 mL of absolute ethanol in an HPLC glass vial ready for analysis.

[0247] Method 4 HPLC sample preparation of final product sample: 1mL of absolute ethanol is added directly to 1 mg of sample in an HPLC glass vial. Ensure full dissolution before submitting to HPLC for analysis.

[0248] The HPLC parameters used for determination of percent diastereomeric purity of product according to Method 4 are shown in Table 7.

Table 7

Instrument	Agilent HPLC 1200 series (Normal Phase HPLC)					
Column	Chiralpak AD-H, 250 x 4.6 mm, 5 μm					
Mobile Phase	A: Heptane, B: Ethanol absolute (Gradient)					
		Time(min)	%A	%B	Flow Rate (mL/min)	
		0.0	80	20	1.20	
		15.5	80	20	1.20	
		28.0	50	50	1.00	
		31.0	50	50	1.00	
		35.0	80	20	1.20	
		40.0	80	20	1.20	
Detection Wavelength	230.0 nm					
Column Temperature	35 oC					
Injection Volume	5 μL					
Run time	40.0 min					
Retention times	Product [Compound (1)]: 12.75 min (R,R,S) diastereomer: 11.99 min					

SEQUENCE LISTING

[0249]

5 <110> CROWE, Michael A.
 ALVIZO, Oscar
 BEHROUZIAN, Behnaz
 BONG, Yong Koy
 COLLIER, Steven
 10 GOHEL, Anupam
 MAVINAHALLI, Jagadeesh
 MODUKURU, Naga
 MUNDORFF, Emily
 SMITH, Derek
 15 SONG, Shiwei
 YEO, Wan Lin

 <120> BIOCATALYSTS FOR EZETIMIBE SYNTHESIS

20 <130> CX2-037USP1

 <160> 168

 <170> PatentIn version 3.5

25 <210> 1
 <211> 759
 <212> DNA
 <213> Artificial Sequence

30 <220>
 <223> Variant of L. kefir ketoreductase

 <400> 1

35

40

45

50

55

EP 2 566 497 B1

	atgaccaacc gtctgaagag caaagtagcc atcgtaaccg gcgggaccca gggatatcgg	60
	ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac cggtcgccac	120
5	gcggatgtag gtgaaaaggc cgccaaatca atcggcggta ctgatgttat tcgctttgtc	180
	cagcaogatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca	240
	ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg cactgcagaa aagcgttgaa	300
10	gacactacca cggaggaatg gcgcaaactg ttgtccgtta atctggatag tgtttttttc	360
	ggcaccogtc tgggcattcg ccgcatgaag aataaaggct tgggcgctag catcatcaat	420
15	atgagcagta tcagtgggct ggtaggcgat ccgatggagg gggcatacaa tgcttccaag	480
	ggggcggtag gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat	540
	gtgcgtgtca acacagtaca tccgggctgt atcaagaccc cgctggtcga taaagatcca	600
20	ggtgctgagg aaatgatgtc acagcgtacg agaacccta tgggccacat tggcgaaccg	660
	aatgacgtgg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt	720
	agcgaatttg tggtcgacgg cgggtatacc gcacagtga	759
25	<210> 2	
	<211> 252	
	<212> PRT	
	<213> Artificial Sequence	
30	<220>	
	<223> Variant of L. kefir ketoreductase	
	<400> 2	
35		
40		
45		
50		
55		

EP 2 566 497 B1

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

245

250

5 <210> 3
 <211> 759
 <212> DNA
 <213> Artificial Sequence

10 <220>
 <223> Variant of L. kefir ketoreductase

<400> 3

15	atgaccaacc gtctgaagag caaagtagcc atcgtaaccg gcgggaccca gggatatcggt	60
	ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttatcac cggtcgccgt	120
	gcagatgtag gtgaaaaggc cgccaaatca atcggcggta ctgatgttat tcgctttgtc	180
20	cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca	240
	ttcggcccggt ttacgaccgt cgtgaacaat gcagggattg cactgcagaa aagcgttgaa	300
	gacactacca cggaggaatg gcgcaaactg ttgtccgta atctggatag tgtttttttc	360
25	ggcaccgctc tgggcattcg ccgcatgaag aataaaggct tgggcgctag catcatcaat	420
	atgagcagta tcagtgggct ggtaggcgat ccgatgatcg gggcatacaa tgcttccaag	480
30	ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat	540
	gtgcgtgtca acacagtaca tccgggcgct atcaagacc cgctggtcga taaagatcca	600
	gggtgctgagg taatgatgtc acagcgtacg agaacccta tgggccacat tggcgaaccg	660
35	aatgacgtgg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt	720
	agcgaatttg tggtcgacgg cgggtatacc gcacagtga	759

40 <210> 4
 <211> 252
 <212> PRT
 <213> Artificial Sequence

45 <220>
 <223> Variant of L. kefir ketoreductase

<400> 4

50

55

EP 2 566 497 B1

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

EP 2 566 497 B1

<210> 5
 <211> 759
 <212> DNA
 <213> Artificial Sequence

5

<220>
 <223> Variant of L. kefir ketoreductase

<400> 5

10

atgaccaacc	gtctgaagag	caaagtagcc	atcgtaaccg	gcgggaccca	gggtatcgg	60
ttggcaatcg	ccgataaatt	tgtagaggag	ggtgcgaaag	tagttatcac	cggtcgccgt	120
gcagatgtag	gtgaaaaggc	cgccaaatca	atcggcggt	ctgatgttat	tcgctttgtc	180
cagcacgatg	catccgatga	agcaggctgg	acgaaactgt	tcgacaccac	cgaggaggca	240
ttcggcccgg	ttacgaccgt	cgtgaacaat	gcagggggcca	ctctgcagaa	aagcgttgaa	300
gacactacca	cggaggaatg	gcgcaaactg	ttgtccgtta	atctggatag	tgtttttttc	360
ggcaccgctc	tgggcattcg	ccgcatgaag	aataaaggct	tgggcgctag	catcatcaat	420
atgagcagta	tcagtgggct	ggtaggcgat	ccgatgatcg	gggcatacaa	tgcttccaag	480
ggggcggtac	gtatcatgtc	gaaaagcgca	gcgctggatt	gcgcagtga	ggactacgat	540
gtgcgtgtca	acacagtaca	tccgggcgct	atcaagacct	cgtgactga	taaatttcca	600
ggtgctgagg	aaatgatctc	acagcgtagc	agaaccctta	tgggccacat	tggcgaaccg	660
aatgacgtgg	catggatctg	tgtgtacctg	gcattctgacg	aatcgaaatt	tgcgacgggt	720
agcgaatttg	tggtcgacgg	cgggtatacc	gcacagtga			759

35

<210> 6
 <211> 252
 <212> PRT
 <213> Artificial Sequence

40

<220>
 <223> Variant of L. kefir ketoreductase

<400> 6

45

50

55

EP 2 566 497 B1

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

EP 2 566 497 B1

<210> 7
 <211> 759
 <212> DNA
 <213> Artificial Sequence

5

<220>
 <223> Variant of L. kefir ketoreductase

<400> 7

10

atgaccaacc	gtctgaagag	caaagtagcc	atcgtaaccg	gcgggaccca	gggtatcggt	60
ttggcaatcg	ccgataaatt	tgtagaggag	ggtgcgaaag	tagttatcac	cggtcgccgt	120
gcagatgtag	gtgaaaaggc	cgccaaatca	atcggcggtg	ctgatgttat	tcgctttgtc	180
cagcacgatg	catccgatga	agcaggctgg	acgaaactgt	tcgacaccac	cgaggaggca	240
ttcggcccgg	ttacgaccgt	cgtgaacaat	gcagggattg	cactgcagaa	aagcgttgaa	300
gacactacca	cggaggaatg	gcgcaaactg	ttgtccgtta	atctggatag	tgtttttttc	360
ggcaccgcgc	tgggcattcg	ccgcatgaag	aataaaggct	tgggcgctag	catcatcaat	420
atgagcagta	tcagtgggct	ggtaggcgat	ccgatgggtg	gggcatacaa	tgcttccaag	480
ggggcggtac	gtatcatgtc	gaaaagcgca	gcgctggatt	gcgcagtgaa	ggactacgat	540
gtgcgtgtca	acacagtaca	tccgggctgt	atcaagaccc	cgctggtcga	taaagatcca	600
ggtgctgagg	aaatgatgtc	acagcgtacg	agaaccctta	tgggccacat	tggcgaaccg	660
aatgacgtgg	catggatctg	tgtgtacctg	gcctctgacg	aatcgaaatt	tgcgacgggt	720
agcgaatttg	tggtcgacgg	cgggtatacc	gcacagtga			759

35

<210> 8
 <211> 252
 <212> PRT
 <213> Artificial Sequence

40

<220>
 <223> Variant of L. kefir ketoreductase

<400> 8

45

50

55

EP 2 566 497 B1

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

EP 2 566 497 B1

<210> 9
 <211> 759
 <212> DNA
 <213> Artificial Sequence

5

<220>
 <223> Variant of L. kefir ketoreductase

<400> 9

10

```

atgaccaacc gtctgaagag caaagtagcc atcgtaaccg gcgggaccca gggatatcgg      60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttatcac cggtcgcgct      120
gcagatgtag gtgaaaaggc cgccaaatca atcggcggta ctgatgttat tcgctttgtc      180
cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca      240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg cactgcagaa aagcgttgaa      300
gacactacca cggaggaatg gcgcaaactg ttgtccgta atctggatag tgtttttttc      360
ggcaccgcgc tgggcattcg ccgcatgaag aataaaggct tgggcgctag catcatcaat      420
atgagcagta tcagtgggct gataggcgat ccgatggtgg gggcatacaa tgcttccaag      480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat      540
gtgcgtgtca acacagtaca tccgggctgt atcaagacc cgctggtcga taaagatcca      600
gggtgtgagg aatgatgtc acagcgtacg agaacccta tgggccacat tggcgaaccg      660
aatgacgtgg catggatctg tgtgtacctg gcactctgac aatcgaaatt tgcgacgggt      720
agcgaatttg tggtcgacgg cgggtatacc gcacagtga      759
  
```

35

<210> 10
 <211> 252
 <212> PRT
 <213> Artificial Sequence

40

<220>
 <223> Variant of L. kefir ketoreductase

<400> 10

45

```

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

Gln Gly Ile Gly Leu Ala Ile Ala Asp Lys Phe Val Glu Glu Gly Ala
                20              25              30

Lys Val Val Ile Thr Gly Arg Arg Ala Asp Val Gly Glu Lys Ala Ala
                35              40              45
  
```

55

EP 2 566 497 B1

Lys Ser Ile Gly Gly Thr Asp Val Ile Arg Phe Val Gln His Asp Ala
50 55 60

5 Ser Asp Glu Ala Gly Trp Thr Lys Leu Phe Asp Thr Thr Glu Glu Ala
65 70 75 80

10 Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ile Ala Leu Gln
85 90 95

Lys Ser Val Glu Asp Thr Thr Thr Glu Glu Trp Arg Lys Leu Leu Ser
100 105 110

15 Val Asn Leu Asp Ser Val Phe Phe Gly Thr Arg Leu Gly Ile Arg Arg
115 120 125

20 Met Lys Asn Lys Gly Leu Gly Ala Ser Ile Ile Asn Met Ser Ser Ile
130 135 140

25 Ser Gly Leu Ile Gly Asp Pro Met Val Gly Ala Tyr Asn Ala Ser Lys
145 150 155 160

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

30 Lys Asp Tyr Asp Val Arg Val Asn Thr Val His Pro Gly Cys Ile Lys
180 185 190

35 Thr Pro Leu Val Asp Lys Asp Pro Gly Ala Glu Glu Met Met Ser Gln
195 200 205

40 Arg Thr Arg Thr Pro Met Gly His Ile Gly Glu Pro Asn Asp Val Ala
210 215 220

Trp Ile Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly
225 230 235 240

45 Ser Glu Phe Val Val Asp Gly Gly Tyr Thr Ala Gln
245 250

<210> 11

50 <211> 759

<212> DNA

<213> Artificial Sequence

<220>

55 <223> Variant of L. kefir ketoreductase

<400> 11

atgaccaacc gtctgaagag caaagtagcc atcgtaacgg gcgggaccca gggatatcgg 60
 ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttatcac cggtcgccgt 120
 5 gcagatgtag gtgaaaaggc cgccaaatca atcggcggta ctgatgttat tcgctttgtc 180
 cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
 10 ttcgggcccg ttacgaccgt cgtgaacaat gcagggattg cactgcagaa aagcgttgaa 300
 gacactacca cggaggaatg gcgcaaactg ttgtccgtta atctggatag tgtttttttc 360
 ggcacccgtc tgggcattcg ccgcatgaag aataaaggct tgggcgctag catcatcaat 420
 15 atgagcagta tcagtgggct ggtaggcgat ccgatggagg gggcatacaa tgcttccaag 480
 ggggcgggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat 540
 20 gtgcgtgtca acacagtaca tccgggctgt atcaagaccc cgctggccga taaagatcca 600
 ggtgctgagg aaatgatgtc acagcgtacg agaacccta tgggccacat tggcgaaccg 660
 aatgacgtgg catggatctg tgtgtacctg gcactctgacg aatcgaaatt tgcgacgggt 720
 25 agcgaatttg tggtcgacgg cgggtatacc gcacagtga 759

<210> 12

<211> 252

<212> PRT

30 <213> Artificial Sequence

<220>

<223> Variant of L. kefir ketoreductase

35 <400> 12

40

45

50

55

EP 2 566 497 B1

	Met	Thr	Asn	Arg	Leu	Lys	Ser	Lys	Val	Ala	Ile	Val	Thr	Gly	Gly	Thr
	1				5					10					15	
5	Gln	Gly	Ile	Gly	Leu	Ala	Ile	Ala	Asp	Lys	Phe	Val	Glu	Glu	Gly	Ala
				20					25					30		
10	Lys	Val	Val	Ile	Thr	Gly	Arg	Arg	Ala	Asp	Val	Gly	Glu	Lys	Ala	Ala
			35					40					45			
15	Lys	Ser	Ile	Gly	Gly	Thr	Asp	Val	Ile	Arg	Phe	Val	Gln	His	Asp	Ala
		50					55					60				
20	Ser	Asp	Glu	Ala	Gly	Trp	Thr	Lys	Leu	Phe	Asp	Thr	Thr	Glu	Glu	Ala
	65					70					75					80
25	Phe	Gly	Pro	Val	Thr	Thr	Val	Val	Asn	Asn	Ala	Gly	Ile	Ala	Leu	Gln
					85					90					95	
30	Lys	Ser	Val	Glu	Asp	Thr	Thr	Thr	Glu	Glu	Trp	Arg	Lys	Leu	Leu	Ser
				100					105					110		
35	Val	Asn	Leu	Asp	Ser	Val	Phe	Phe	Gly	Thr	Arg	Leu	Gly	Ile	Arg	Arg
			115					120					125			

EP 2 566 497 B1

Met Lys Asn Lys Gly Leu Gly Ala Ser Ile Ile Asn Met Ser Ser Ile
130 135 140

5 Ser Gly Leu Val Gly Asp Pro Met Val Gly Ala Tyr Asn Ala Ser Lys
145 150 155 160

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

Lys Asp Tyr Asp Val Arg Val Asn Thr Val His Pro Gly Cys Ile Lys
180 185 190

15 Thr Pro Leu Ala Asp Lys Asp Pro Gly Ala Glu Glu Met Met Ser Gln
195 200 205

20 Arg Thr Arg Thr Pro Met Gly His Ile Gly Glu Pro Asn Asp Val Ala
210 215 220

25 Trp Ile Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly
225 230 235 240

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

30

<210> 13

<211> 759

<212> DNA

<213> Artificial Sequence

35

<220>

<223> Variant of L. kefir ketoreductase

40

<400> 13

45

50

55

EP 2 566 497 B1

	atgaccaacc gtctgaagag caaagtagcc atcgtaaccg gcgggaccca gggatatcgg	60
	ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttatcac cggtcgccgt	120
5	gcagatgtag gtgaaaaggc cgccaaatca atcggcggta ctgatgttat tcgctttgtc	180
	cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca	240
	ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg cactgcagaa aagcgttgaa	300
10	gacactacca cggaggaatg gcgcaaactg ttgtccgtta atctggatag tggttttttc	360
	ggcaccgcgc tgggcattcg ccgcatgaag aataaaggct tgggcgctag catcatcaat	420
	atgagcagta tcagtgggct ggtaggcgat ccgatggtgg gggcatacaa tgcttccaag	480
15	ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat	540
	gtgcgtgtca acacagtaca tccgggctgt atcaagacc cgctggtcga taaagatcca	600
20	ggtgctgagg aaatgatgac acagcgtacg agaacccta tgggccacat tggcgaaccg	660
	aatgacgtgg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt	720
25	agcgaatttg tggtcgacgg cgggtatacc gcacagtga	759

	<210> 14
	<211> 252
	<212> PRT
30	<213> Artificial Sequence
	<220>
	<223> Variant of L. kefir ketoreductase
35	<400> 14

40

45

50

55

EP 2 566 497 B1

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

EP 2 566 497 B1

<210> 15
<211> 759
<212> DNA
<213> Artificial Sequence

5

<220>
<223> Variant of L. kefir ketoreductase

<400> 15

10

```

atgaccaacc gtctgaagag caaagtagcc atcgtaaccg gcgggaccca gggatatcgg      60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttatcac cggtcgccgt      120
gcagatgtag gtgaaaaggc cgccaaatca atcggcggta ctgatgttat tcgctttgtc      180
cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca      240
ttcggccccg ttacgaccgt cgtgaacaat gcagggattg cactggttaa aagcgttgaa      300
gacactacca cggaggaatg gcgcaaactg ttgtccgta atctggatag tgtttttttc      360
ggcaccctgc tgggcattcg ccgcatgaag aataaaggct tgggcgctag catcatcaat      420
atgagcagta tcagtgggct ggtaggcgat ccgatggtgg gggcatacaa tgcttccaag      480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat      540
gtgcgtgtca acacagtaca tccgggctgt atcaagacc cgctggtcga taaagatcca      600
gggtgtgagg aaatgatgtc acagcgtacg agaaccctta tgggccacat tggcgaaccg      660
aatgacgtgg catggatctg tgtgtacctg gcactctgac aatcgaaatt tgcgacgggt      720
agcgaatttg tggtcgacgg cgggtatacc gcacagtga      759

```

35

<210> 16
<211> 252
<212> PRT
<213> Artificial Sequence

40

<220>
<223> Variant of L. kefir ketoreductase

<400> 16

45

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

50

Gln Gly Ile Gly Leu Ala Ile Ala Asp Lys Phe Val Glu Glu Gly Ala
20 25 30

55

EP 2 566 497 B1

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

<210> 17

<211> 759

55 <212> DNA

<213> Artificial Sequence

<220>

EP 2 566 497 B1

<223> Variant of L. kefir ketoreductase

<400> 17

5	atgaccaacc gtctgaagag caaagtagcc atcgtaaccg gcgggaccca gggatatcgg	60
	ttggcaatcg ccgataaatt thtagaggag ggtgcgaaag tagttatcac cggtcgccgt	120
10	gcagatgtag gtgaaaaggc cgccaaatca atcggcggta ctgatgttat tcgctttgtc	180
	cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca	240
	ttcggcccggtt ttacgaccgt cgtgaacaat gcagggattg cactgcagaa aagcgttgaa	300
15	gacactacca cggaggaatg gcgcaaactg ttgtccgtta atctggatag tgtttttttc	360
	ggcaccgcgtc tgggcattcg ccgcatgaag aataaaggct tgggcgctag catcatcaat	420
	atgagcagta tcagtgggct ggtaggcgat ccgatggtgg gggcatacaa tgcttccaag	480
20	ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat	540
	gtgcgtgtca acacagtaca tccgggctgt atcaagacc cgctgtctga taaagatcca	600
	ggtgctgagg aaatgatgtc acagcgtacg agaacccta tgggccacat tggcgaaccg	660
25	aatgacgtgg catggatctg tgtgtacctg gcattctgacg aatcgaaatt tgcgacgggt	720
	agcgaatttg tggtcgacgg cgggtatacc gcacagtga	759

30 <210> 18
 <211> 252
 <212> PRT
 <213> Artificial Sequence

35 <220>
 <223> Variant of L. kefir ketoreductase
 <400> 18

40

45

50

55

EP 2 566 497 B1

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

EP 2 566 497 B1

<210> 19
 <211> 759
 <212> DNA
 <213> Artificial Sequence

5

<220>
 <223> Variant of L. kefir ketoreductase

<400> 19

10

atgaccaacc	gtctgaagag	caaagtagcc	atcgtaaccg	gcgggaccca	gggtatcggt	60
ttggcaatcg	ccgataaatt	tgtagaggag	ggcgcgaaag	tagttatcac	cggtcgccgt	120
gcagatgtag	gtgaaaaggc	cgccaaatca	atcggcggta	ctgatgttat	tcgctttgtc	180
cagcacgatg	catccgatga	agcaggctgg	acgaaactgt	tcgacaccac	cgaggaggca	240
ttcggcccgg	ttacgaccgt	cgtgaacaat	gcagggattg	cactgcagaa	aagcgttgaa	300
gacactacca	cggaggaatg	gcgcaaactg	ttgtccgtta	atctggatag	tgtttttttc	360
ggcaccocgtc	tgggcattcg	ccgcatgaag	aataaaggct	tgggcgctag	catcatcaat	420
atgagcagta	tcagtgggct	ggtaggcgat	ccgatggtgg	gggcatacaa	tgcttccaag	480
ggggcggtac	gtatcatgtc	gaaaagcgca	gcgctggatt	gcgcagtga	ggactacgat	540
gtgcgtgtca	acacagtaca	tccgggctgt	atcaagaccc	cgctgtgtga	taaagatcca	600
gggtctgagg	aaatgatgtc	acagcgtacg	agaacccta	tgggccacat	tggcgaaccg	660
aatgacgtgg	catggatctg	tgtgtacctg	gcacctgacg	aatcgaaatt	tgcgacgggt	720
agcgaatttg	tggtcgacgg	cgggtatacc	gcacagtga			759

35

<210> 20
 <211> 252
 <212> PRT
 <213> Artificial Sequence

40

<220>
 <223> Variant of L. kefir ketoreductase

<400> 20

45

50

55

EP 2 566 497 B1

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

EP 2 566 497 B1

<210> 21
 <211> 759
 <212> DNA
 <213> Artificial Sequence

5

<220>
 <223> Variant of L. kefir ketoreductase

<400> 21

10

atgaccaacc	gtctgaagag	caaagtagcc	atcgtaaccg	gcgggaccca	gggtatcgg	60
ttggcaatcg	ccgataaatt	tgtagaggag	ggtgcgaaag	tagttatcac	cggtcgcg	120
gcagatgtag	gtgaaaaggc	cgccaaatca	atcggcggt	ctgatgttat	tcgctttgtc	180
cagcacgatg	catccgatga	agcaggettg	acgaaactgt	tcgacaccac	cgaggaggca	240
ttcggccccg	ttacgaccgt	cgtgaacaat	gcagggattg	cactgcagaa	aagcgttgaa	300
gacactacca	cggaggaatg	gcgcaaactg	ttgtccgtta	atctggatag	tgtttttttc	360
ggcaccgcgc	tgggcattcg	ccgcatgaag	aataaaggct	tgggcgctag	catcatcaat	420
atgagcagta	tcagtgggct	ggtaggcgat	ccgatggtgg	gggcatacaa	tgcttccaag	480
ggggcggtac	gtatcatgtc	gaaaagcgca	gcgctggatt	gcgcagtgaa	ggactacgat	540
gtgcgtgtca	acacagtaca	tccgggctgt	atcaagaccc	cgctgaatga	taaagatcca	600
ggtgctgagg	aaatgatgtc	acagcgtacg	agaaccctta	tgggccacat	tggcgaaccg	660
aatgacgtgg	catggatctg	tgtgtacctg	gcacctgacg	aatcgaaatt	tgcgacgggt	720
agcgaatttg	tggtcgacgg	cgggtatacc	gcacagtga			759

35

<210> 22
 <211> 252
 <212> PRT
 <213> Artificial Sequence

40

<220>
 <223> Variant of L. kefir ketoreductase

<400> 22

45

Met Thr Asn Arg Leu Lys Ser Lys Val Ala Ile Val Thr Gly Gly Thr

50

55

EP 2 566 497 B1

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

EP 2 566 497 B1

<210> 23
 <211> 759
 <212> DNA
 <213> Artificial Sequence

5

<220>
 <223> Variant of L. kefir ketoreductase

<400> 23

10

```

atgaccaacc gtctgaagag caaagtagcc atcgtaaccg gcgggaccca gggatatcgg      60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttatcac cggtcgcgct      120
gcagatgtag gtgaaaaggc cgccaaatca atcggcggtta ctgatgttat tcgctttgtc      180
cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca      240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg cactgcagaa aagcgttgaa      300
gacactacca cggaggaatg gcgcaaactg ttgtccgta atctggatag tgtttttttc      360
ggcaccgctc tgggcattcg ccgcatgaag aataaaggct tgggcgctag catcatcaat      420
atgagcagta tcagtgggct ggtaggcgat ccgatggtgg gggcatacaa tgcttccaag      480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat      540
gtgcgtgtca acacagtaca tccgggctgt atcaagacc cgctggtcga taaagatcca      600
ggtttgagg aaatgatgtc acagcgtacg agaacccta tgggccacat tggcgaaccg      660
aatgacgtgg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt      720
agcgaatttg tggtcgacgg cgggtatacc gcacagtga      759
  
```

35

<210> 24
 <211> 252
 <212> PRT
 <213> Artificial Sequence

40

<220>
 <223> Variant of L. kefir ketoreductase

<400> 24

45

50

55

EP 2 566 497 B1

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

25

30

35

40

45

50

55

EP 2 566 497 B1

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

<210> 25

<211> 759

<212> DNA

<213> Artificial Sequence

<220>

<223> Variant of L. kefir ketoreductase

<400> 25

EP 2 566 497 B1

	atgaccaacc gtctgaagag caaagtagcc atcgtaaccg gcgggaccca gggatatcgg	60
	ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttatcac cggtcgccgt	120
5	gcagatgtag gtgaaaaggc cgccaaatca atcggcggta ctgatgttat tcgctttgtc	180
	cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca	240
	ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg cactgcagaa aagcgttgaa	300
10	gacactacca cggaggaatg gcgcaaactg ttgtccgtta atctggatag tgtttttttc	360
	ggcaccocgtc tgggcattcg ccgcatgaag aataaaggct tgggcgctag catcatcaat	420
15	atgagcagta tcagtgggct ggtaggcgat ccgatggtag gggcatacaa tgcttccaag	480
	ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat	540
	gtgcgtgtca acacagtaca tccgggctgt atcaagaccc cgctggtcga taaagatcca	600
20	ggtaatgagg aaatgatgtc acagcgtacg agaacccta tgggccacat tggcgaaccg	660
	aatgacgtgg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt	720
25	agcgaatttg tggtcgacgg cgggtatacc gcacagtga	759

<210> 26

<211> 252

<212> PRT

30 <213> Artificial Sequence

<220>

<223> Variant of L. kefir ketoreductase

35 <400> 26

40

45

50

55

EP 2 566 497 B1

	Met	Thr	Asn	Arg	Leu	Lys	Ser	Lys	Val	Ala	Ile	Val	Thr	Gly	Gly	Thr	
	1				5					10					15		
5	Gln	Gly	Ile	Gly	Leu	Ala	Ile	Ala	Asp	Lys	Phe	Val	Glu	Glu	Gly	Ala	
				20					25					30			
10	Lys	Val	Val	Ile	Thr	Gly	Arg	Arg	Ala	Asp	Val	Gly	Glu	Lys	Ala	Ala	
			35					40					45				
15	Lys	Ser	Ile	Gly	Gly	Thr	Asp	Val	Ile	Arg	Phe	Val	Gln	His	Asp	Ala	
		50					55					60					
20	Ser	Asp	Glu	Ala	Gly	Trp	Thr	Lys	Leu	Phe	Asp	Thr	Thr	Glu	Glu	Ala	
	65					70					75					80	
25	Phe	Gly	Pro	Val	Thr	Thr	Val	Val	Asn	Asn	Ala	Gly	Ile	Ala	Leu	Gln	
					85					90					95		
30	Lys	Ser	Val	Glu	Asp	Thr	Thr	Thr	Glu	Glu	Trp	Arg	Lys	Leu	Leu	Ser	
				100					105					110			
35	Val	Asn	Leu	Asp	Ser	Val	Phe	Phe	Gly	Thr	Arg	Leu	Gly	Ile	Arg	Arg	
			115					120					125				
40	Met	Lys	Asn	Lys	Gly	Leu	Gly	Ala	Ser	Ile	Ile	Asn	Met	Ser	Ser	Ile	
		130					135					140					
45	Ser	Gly	Leu	Val	Gly	Asp	Pro	Met	Val	Gly	Ala	Tyr	Asn	Ala	Ser	Lys	
	145					150					155					160	
50	Gly	Ala	Val	Arg	Ile	Met	Ser	Lys	Ser	Ala	Ala	Leu	Asp	Cys	Ala	Val	
55																	

EP 2 566 497 B1

[illegible]

EP 2 566 497 B1

<212> PRT

<213> Artificial Sequence

<220>

5 <223> Variant of L. kefir ketoreductase

<400> 28

10

15

20

25

30

35

40

45

50

55

EP 2 566 497 B1

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

245

250

5 <210> 29
 <211> 759
 <212> DNA
 <213> Artificial Sequence

10 <220>
 <223> Variant of L. kefir ketoreductase

<400> 29

15 atgaccaacc gtctgaagag caaagtagcc atcgtaaccg gcgggaccca gggatatcggt 60
 ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttatcac cggtcgccgt 120
 gcagatgtag gtgaaaaggc cgccaaatca atcggcggtta ctgatgttat tcgctttgtc 180
 20 cagcacgatg catccgatga agcaggettg acgaaactgt tcgacaccac cgaggaggca 240
 ttccggcccg ttacgaccgt cgtgaacaat gcagggggccc ctctgcagaa aagcgttgaa 300
 gacactacca cggaggaatg gcgcaaactg ttgtccgtta atctggatag tgtttttttc 360
 25 ggcacccgtc tgggcattcg ccgcatgaag aataaaggct tgggcgctag catcatcaat 420
 atgagcagta tcagtgggct ggtaggcgat ccgatgatcg gggcatacaa tgcttccaag 480
 30 ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat 540
 gtgcgtgtca acacagtaca tccgggcgct atcaagacc cgctgactga taaatttcca 600
 ggtgctgagg aaatgatctc acagcgtacg agaacccta tgggccacat tggcgaaccg 660
 35 aatgacgtgg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
 agcgaatttg tggtcgacgg cgggtatacc gcacagtga 759

40 <210> 30
 <211> 252
 <212> PRT
 <213> Artificial Sequence

45 <220>
 <223> Variant of L. kefir ketoreductase

<400> 30

50

55

EP 2 566 497 B1

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

EP 2 566 497 B1

<210> 31
 <211> 759
 <212> DNA
 <213> Artificial Sequence

5

<220>
 <223> Variant of L. kefir ketoreductase

<400> 31

10

atgaccaacc	gtctgaagag	caaagtagcc	atcgtaaccg	gcgggaccca	gggtatcggt	60
ttggcaatcg	ccgataaatt	tgtagaggag	ggtgcgaaag	tagttatcac	cggtcgccgt	120
gcagatgtag	gtgaaaaggc	cgccaaatca	atcggcggtg	ctgatgttat	tcgctttgtc	180
cagcacgatg	catccgatga	agcaggctgg	acgaaactgt	tcgacaccac	cgaggaggca	240
ttcggcccgg	ttacgaccgt	cgtgaacaat	gcagggattg	cactgggcaa	aagcgttgaa	300
gacactacca	cggaggaatg	gcgcaaactg	ttgtccgtta	atctggatag	tgtttttttc	360
ggcacccgtc	tgggcattcg	ccgcatgaag	aataaaggct	tgggcgctag	catcatcaat	420
atgagcagta	tcagtgggct	ggtaggcgat	ccgatgatcg	gggcatacaa	tgcttccaag	480
ggggcggtac	gtatcatgtc	gaaaagcgca	gcgctggatt	gcgcagtga	ggactacgat	540
gtgctgttca	acacagtaca	tccgggcgct	atcaagacct	cgctgactga	taaattttcca	600
ggtgctgagg	aaatgatctc	acagcgtacg	agaaccctta	tgggccacat	tggcgaaccg	660
aatgacgtgg	catggatctg	tgtgtacctg	gcctctgacg	aatcgaaatt	tgcgacgggt	720
agcgaatttg	tggtcgacgg	cgggtatacc	gcacagtga			759

35

<210> 32
 <211> 252
 <212> PRT
 <213> Artificial Sequence

40

<220>
 <223> Variant of L. kefir ketoreductase

<400> 32

45

50

55

EP 2 566 497 B1

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

EP 2 566 497 B1

<210> 33
 <211> 759
 <212> DNA
 <213> Artificial Sequence

5

<220>
 <223> Variant of L. kefir ketoreductase

<400> 33

10

```

atgaccaacc gtctgaagag caaagtagcc atcgtaaccc gcgggaccca gggatatcgg      60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttatcac cggtcgccgt      120
gcagatgtag gtgaaaaggc cgccaaatca atcggcggtta ctgatgttat tcgctttgtc      180
cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca      240
ttcggccccg ttacgaccgt cgtgaacaat gcagggattg cactggctaa aagcgttgaa      300
gacactacca cggaggaatg gcgcaaactg ttgtccgta atctggatag tgtttttttc      360
ggcaccctgc tgggcattcg ccgcatgaag aataaaggct tgggcgctag catcatcaat      420
atgagcagta tcagtgggct ggtaggcgat ccgatgatcg gggcatacaa tgcttccaag      480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat      540
gtgcgtgtca acacagtaca tccgggcgct atcaagacc cgctgactga taaatttcca      600
gggtgctgagg aaatgatctc acagcgtacg agaacccta tgggccacat tggcgaaccg      660
aatgacgtgg catggatctg tgtgtacctg gcctctgacg aatcgaaatt tgcgacgggt      720
agcgaatttg tggtcgacgg cgggtatacc gcacagtga      759
  
```

35

<210> 34
 <211> 252
 <212> PRT
 <213> Artificial Sequence

40

<220>
 <223> Variant of L. kefir ketoreductase

<400> 34

45

50

55

EP 2 566 497 B1

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

EP 2 566 497 B1

<210> 35
 <211> 759
 <212> DNA
 <213> Artificial Sequence

5

<220>
 <223> Variant of L. kefir ketoreductase

<400> 35

10

```

atgaccaacc gtctgaagag caaagtagcc atcgtaaccg gcgggaccca gggatatcggt      60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttatcac cggtcgccgt      120
gcagatgtag gtgaaaaggc cgccaaatca atcggcggta ctgatgttat tcgctttgtc      180
cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca      240
ttcggccccg ttacgaccgt cgtgaacaat gcagggggcca ctctgcagaa aagcgttgaa      300
gacactacca cggaggaatg gcgcaaactg ttgtccgtta atctggatag tgtttttttc      360
ggcaccctgc tgggcattcg ccgcatgaag aataaaggct tgggcgctag catcatcaat      420
atgagctcga tcagtgggct ggtaggcgat ccgatgatcg gggcatacaa tgcttccaag      480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat      540
gtgcgtgtca acacagtaca tccgggcgct atcaagacct cgctgactga taaatttcca      600
gggtgctgggg aaatgatctc acagcgtacg agaaccctta tgggccacat tggcgaaccg      660
aatgacgtgg catggatctg tgtgtacctg gcacctgacg aatcgaaatt tgcgacgggt      720
agcgaatttg tggtcgacgg cgggtatacc gcacagtga      759
  
```

35

<210> 36
 <211> 252
 <212> PRT
 <213> Artificial Sequence

40

<220>
 <223> Variant of L. kefir ketoreductase

<400> 36

45

```

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

Gln Gly Ile Gly Leu Ala Ile Ala Asp Lys Phe Val Glu Glu Gly Ala
20           25           30

Lys Val Val Ile Thr Gly Arg Arg Ala Asp Val Gly Glu Lys Ala Ala
35           40           45
  
```

55

EP 2 566 497 B1

Lys Ser Ile Gly Gly Thr Asp Val Ile Arg Phe Val Gln His Asp Ala
50 55 60

5 Ser Asp Glu Ala Gly Trp Thr Lys Leu Phe Asp Thr Thr Glu Glu Ala
65 70 75 80

10 Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ala Thr Leu Gln
85 90 95

15 Lys Ser Val Glu Asp Thr Thr Thr Glu Glu Trp Arg Lys Leu Leu Ser
100 105 110

Val Asn Leu Asp Ser Val Phe Phe Gly Thr Arg Leu Gly Ile Arg Arg
115 120 125

20 Met Lys Asn Lys Gly Leu Gly Ala Ser Ile Ile Asn Met Ser Ser Ile
130 135 140

25 Ser Gly Leu Val Gly Asp Pro Met Ile Gly Ala Tyr Asn Ala Ser Lys
145 150 155 160

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

30 Lys Asp Tyr Asp Val Arg Val Asn Thr Val His Pro Gly Ala Ile Lys
180 185 190

35 Thr Pro Leu Thr Asp Lys Phe Pro Gly Ala Gly Glu Met Ile Ser Gln
195 200 205

40 Arg Thr Arg Thr Pro Met Gly His Ile Gly Glu Pro Asn Asp Val Ala
210 215 220

Trp Ile Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly
225 230 235 240

45 Ser Glu Phe Val Val Asp Gly Gly Tyr Thr Ala Gln
245 250

<210> 37

50 <211> 759

<212> DNA

<213> Artificial Sequence

<220>

55 <223> Variant of L. kefir ketoreductase

<400> 37

	atgaccaacc gtctgaagag caaagtagcc atcgtaaccg gcgggaccca gggatatcgg	60
	ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttatcac cggtcgccgt	120
5	gcagatgtag gtgaaaaggc cgccaaatca atcggcggtta ctgatgttat tcgctttgtc	180
	cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca	240
10	ttcggcccgg ttacgaccgt cgtgaacaat gcaggggcca ctctgggtgaa aagcgttgaa	300
	gacactacca cggaggaatg gcgcaaactg ttgtccgtta atctggatag tgtttttttc	360
	ggcaccgctc tgggcattcg ccgcatgaag aataaaggct tgggcgctag catcatcaat	420
15	atgagcagta tcagtgggct ggtaggcgat ccgatgatcg gggcatacaa tgcttccaag	480
	ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtgaa ggactacgat	540
20	gtgcgtgtca acacagtaca tccgggcgct atcaagacct cgctgaccga taaatttcca	600
	gggtgctgagg aaatgatctc acagcgtacg agaacccta tgggccacat tggcgaaccg	660
	aatgacgtgg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt	720
25	agcgaatttg tggtcgacgg cgggtatacc gcacagtga	759

<210> 38

<211> 252

<212> PRT

30 <213> Artificial Sequence

<220>

<223> Variant of L. kefir ketoreductase

35 <400> 38

40

45

50

55

EP 2 566 497 B1

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

5 Gln Gly Ile Gly Leu Ala Ile Ala Asp Lys Phe Val Glu Glu Gly Ala
20 25 30

10 Lys Val Val Ile Thr Gly Arg Arg Ala Asp Val Gly Glu Lys Ala Ala
35 40 45

15 Lys Ser Ile Gly Gly Thr Asp Val Ile Arg Phe Val Gln His Asp Ala
50 55 60

Ser Asp Glu Ala Gly Trp Thr Lys Leu Phe Asp Thr Thr Glu Glu Ala
65 70 75 80

20 Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ala Thr Leu Val
85 90 95

25 Lys Ser Val Glu Asp Thr Thr Thr Glu Glu Trp Arg Lys Leu Leu Ser
100 105 110

30 Val Asn Leu Asp Ser Val Phe Phe Gly Thr Arg Leu Gly Ile Arg Arg
115 120 125

35

40

45

50

55

EP 2 566 497 B1

	Met	Lys	Asn	Lys	Gly	Leu	Gly	Ala	Ser	Ile	Ile	Asn	Met	Ser	Ser	Ile
	130						135					140				
5	Ser	Gly	Leu	Val	Gly	Asp	Pro	Met	Ile	Gly	Ala	Tyr	Asn	Ala	Ser	Lys
	145					150					155					160
10	Gly	Ala	Val	Arg	Ile	Met	Ser	Lys	Ser	Ala	Ala	Leu	Asp	Cys	Ala	Val
					165					170					175	
15	Lys	Asp	Tyr	Asp	Val	Arg	Val	Asn	Thr	Val	His	Pro	Gly	Ala	Ile	Lys
				180					185					190		
20	Thr	Pro	Leu	Thr	Asp	Lys	Phe	Pro	Gly	Ala	Glu	Glu	Met	Ile	Ser	Gln
			195					200					205			
25	Arg	Thr	Arg	Thr	Pro	Met	Gly	His	Ile	Gly	Glu	Pro	Asn	Asp	Val	Ala
	210						215					220				
30	Trp	Ile	Cys	Val	Tyr	Leu	Ala	Ser	Asp	Glu	Ser	Lys	Phe	Ala	Thr	Gly
	225					230					235					240
35	Ser	Glu	Phe	Val	Val	Asp	Gly	Gly	Tyr	Thr	Ala	Gln				
					245					250						

<210> 39

<211> 759

<212> DNA

<213> Artificial Sequence

<220>

<223> Variant of L. kefir ketoreductase

<400> 39

EP 2 566 497 B1

	atgaccaacc gtctgaagag caaagtagcc atcgtaaccg gcgggaccca gggatatcgg	60
	ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttatcac cggtcgccgt	120
5	gcagatgtag gtgaaaaggc cgccaaatca atcggcggtta ctgatgttat tcgctttgtc	180
	cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca	240
	ttcggccccg ttacgaccgt cgtgaacaat gcaggggcca ctctgcagaa aagcgttgaa	300
10	gacactacca cggaggaatg gcgcaaactg ttgtccgtta atctggatag tgtttttttc	360
	ggcaccctgc tgggcattcg ccgcatgaag aataaaggct tgggcgctag catcatcaat	420
	atgagcagta tcagtgggct ggtaggcgat ccgatgatcg gggcatacaa tgcttccaag	480
15	ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat	540
	gtgcgtgtca acacagtaca tccgggcgct atcaagacc cgatgactga taaatttcca	600
20	ggtgctgagg aaatgatctc acagcgtacg agaacccta tgggccacat tggcgaaccg	660
	aatgacgtgg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt	720
25	agcgaatttg tggtcgacgg cgggtatacc gcacagtga	759

<210> 40

<211> 252

<212> PRT

30 <213> Artificial Sequence

<220>

<223> Variant of L. kefir ketoreductase

35 <400> 40

40

45

50

55

EP 2 566 497 B1

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

EP 2 566 497 B1

<210> 41
 <211> 759
 <212> DNA
 <213> Artificial Sequence

5

<220>
 <223> Variant of L. kefir ketoreductase

<400> 41

10

```

atgaccaacc gtctgaagag caaagtagcc atcgtaaccc gcgggaccca gggatatcgg      60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttatcac cggtcgccgt      120
gcagatgtag gtgaaaaggc cgccaaatca atcggcggta ctgatgttat tcgctttgtc      180
cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca      240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggggcca ctctgcagaa aagcgttgaa      300
gacactacca cggaggaatg gcgcaaactg ttgtccgtta atctggatag tgtttttttc      360
ggcaccgcgc tgggcattcg ccgcatgaag aataaaggct tgggcgctag catcatcaat      420
atgagcagta tcagtgggct ggtaggcgat ccgatgatcg gggcatacaa tgcttccaag      480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat      540
gtgcgtgtca acacagtaca tccgggcgct atcaagacc cgctgactga taaatttcca      600
gggtgtgagg aaatgatctg tcagcgtacg agaacccta tgggccacat tggcgaaccg      660
aatgacgtgg catggatctg tgtgtacctg gcactctgac aatcgaaatt tgcgacgggt      720
agcgaatttg tggtcgacgg cgggtatacc gcacagtga      759
  
```

35

<210> 42
 <211> 252
 <212> PRT
 <213> Artificial Sequence

40

<220>
 <223> Variant of L. kefir ketoreductase

<400> 42

45

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

50

Gln Gly Ile Gly Leu Ala Ile Ala Asp Lys Phe Val Glu Glu Gly Ala
 20 25 30

55

EP 2 566 497 B1

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

<210> 43

<211> 759

55 <212> DNA

<213> Artificial Sequence

<220>

EP 2 566 497 B1

<223> Variant of L. kefir ketoreductase

<400> 43

5	atgaccaacc gtctgaagag caaagtagcc atcgtaaccg gcgggaccca gggatatcgg	60
	ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttatcac cggtcgccgt	120
10	gcagatgtag gtgaaaaggc cgccaaatca atcggcggtta ctgatgttat tcgctttgtc	180
	cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca	240
	ttcggcccggtt ttacgaccgt cgtgaacaat gcagggggcca ctctgcagaa aagcgttgaa	300
15	gacactacca cggaggaatg gcgcaaactg ttgtccgtta atctggatag tgtttttttc	360
	ggcaccgcgtc tgggcattcg ccgcatgaag aataaaggct tgggcgctag catcatcaat	420
	atgagcagta tcagtgggct ggtaggcgat ccgatgatcg gggcatacaa tgcttccaag	480
20	ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat	540
	gtgcgtgtca acacagtaca tccgggcgct atcaagacc cgctgactga taaatttcca	600
	ggtgctgagg aaatgatcat tcagcgtacg agaacccta tgggccacat tggcgaaccg	660
25	aatgacgtgg catggatctg tgtgtacctg gcattctgacg aatcgaaatt tgcgacgggt	720
	agcgaatttg tggtcgacgg cgggtatacc gcacagtga	759

30 <210> 44
 <211> 252
 <212> PRT
 <213> Artificial Sequence

35 <220>
 <223> Variant of L. kefir ketoreductase
 <400> 44

40

45

50

55

EP 2 566 497 B1

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

EP 2 566 497 B1

<210> 45
 <211> 759
 <212> DNA
 <213> Artificial Sequence

5

<220>
 <223> Variant of L. kefir ketoreductase

<400> 45

10

```

atgaccaacc gtctgaagag caaagtagcc atcgtaaccg gcgggaccca gggatatcgg      60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttatcac cggtcgccgt      120
gcagatgtag gtgaaaaggc cgccaaatca atcggcggta ctgatgttat tcgctttgtc      180
cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca      240
ttcggcccgg ttacgaccgt cgtgaacaat gcaggggcca ctctgcagaa aagcgttgaa      300
gacactacca cggaggaatg gcgcaaactg ttgtccgtta atctggatag tgtttttttc      360
ggcaccocgtc tgggcattcg ccgcatgaag aataaaggct tgggcgctag catcatcaat      420
atgagcagta tcagtgggct ggtaggcgat ccgatgatcg gggcatacaa tgcttccaag      480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat      540
gtgcgtgtca acacagtaca tccgggcgct atcaagaccc cgctgactga taaattttcca      600
ggtgctgagg aaatgatcaa tcagcgtacg agaacccta tgggccacat tggcgaaccg      660
aatgacgtgg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt      720
agcgaatttg tggtcgacgg cgggtatacc gcacagtga      759
  
```

35

<210> 46
 <211> 252
 <212> PRT
 <213> Artificial Sequence

40

<220>
 <223> Variant of L. kefir ketoreductase

<400> 46

45

50

55

EP 2 566 497 B1

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

EP 2 566 497 B1

<210> 47
 <211> 759
 <212> DNA
 <213> Artificial Sequence

5

<220>
 <223> Variant of L. kefir ketoreductase

<400> 47

10

atgaccaacc	gtctgaagag	caaagtagcc	atcgtaaccg	gcgggaccca	gggtatcggt	60
cgcgcaatcg	cccgtaaatt	tgtagaggag	ggtgcgaaag	tagttatcac	cggtcgccgt	120
gcagatgtag	gtgaaaaggc	cgccaaatca	atcggcggtg	ctgatgttat	ccgctttgtc	180
cagcacgatg	catccgatga	agcaggctgg	acgaaactgt	tcgacaccac	cgaggaggca	240
ttcggcccgg	ttacgaccgt	cgtgaacaat	gcggggggcca	ctctgcagaa	aagcgttgaa	300
gacactacca	cggaggaatg	gcgcaaactg	ttgtccgtta	atctggatgg	ggtttttttc	360
ggcaccgcgc	tgggcattaa	acgcatgaag	aataaaggct	tgggcgctag	catcatcaat	420
atgagctcga	tcagtgggat	ggtaggcgat	ccgatgatcg	gggcatacaa	tgcttccaag	480
ggggcggtac	gtatcatgtc	gaaaagcgca	gcgctggatt	gcgcagtga	ggactacgat	540
gtgcgtgtca	acacagtaca	tccgggcgct	atcaagacc	cgctgactga	taaatttcca	600
ggtgctgggg	aaatgatctc	acagcgtacg	agaaccctta	tgggccacat	tggcgaaccg	660
gacgacgtgg	catggatctg	tgtgtacctg	gcacgtgacg	aatcgaaatt	tgcgacgggt	720
agcgaatttg	tggtcgacgg	cgggtatacc	gcacagtga			759

35

<210> 48
 <211> 252
 <212> PRT
 <213> Artificial Sequence

40

<220>
 <223> Variant of L. kefir ketoreductase

<400> 48

45

Met Thr Asn Arg Leu Lys Ser Lys Val Ala Ile Val Thr Gly Gly Thr

50

55

EP 2 566 497 B1

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

EP 2 566 497 B1

<210> 49
 <211> 759
 <212> DNA
 <213> Artificial Sequence

5

<220>
 <223> Variant of L. kefir ketoreductase

<400> 49

10

```

atgaccaacc gtctgaagag caaagtagcc atcgtaaccg gcgggaccca gggatatcgg      60
tttgcaatcg cccgtaaatt tgtagaggag ggtgcgaaag tagttatcac cggtcgccgt      120
gcagatgtag gtgaaaaggc cgccaaatca atcggcggtta ctgatgttat tcgctttgtc      180
cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca      240
ttcggccccg ttacgaccgt cgtgaacaat gcaggggcca ctctgcagaa aagcgttgaa      300
gacactacca cggaggaatg gcgcaaactg ttgtccgtta atctggatgg ggtttttttc      360
ggcaccctgc tgggcattcg ccgcatgaag aataaaggct tgggcgctag catcatcaat      420
atgagctcga tcagtgggat ggtaggcgat ccgatgatcg gggcatacaa tgcttccaag      480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat      540
gtgcgtgtca acacagtaca tccgggcgct atcaagacc cgctgactga taaatttcca      600
ggtgctgggg aaatgatctc acagcgtacg agaacccta tgggccacat tggcgaaccg      660
gacgacgtgg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt      720
agcgaatttg tggtcgacgg cgggtatacc gcacagtga      759
  
```

35

<210> 50
 <211> 252
 <212> PRT
 <213> Artificial Sequence

40

<220>
 <223> Variant of L. kefir ketoreductase

<400> 50

45

50

55

EP 2 566 497 B1

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

5 Gln Gly Ile Gly Phe Ala Ile Ala Arg Lys Phe Val Glu Glu Gly Ala
20 25 30

10 Lys Val Val Ile Thr Gly Arg Arg Ala Asp Val Gly Glu Lys Ala Ala
35 40 45

15 Lys Ser Ile Gly Gly Thr Asp Val Ile Arg Phe Val Gln His Asp Ala
50 55 60

Ser Asp Glu Ala Gly Trp Thr Lys Leu Phe Asp Thr Thr Glu Glu Ala
65 70 75 80

20 Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ala Thr Leu Gln

25

30

35

40

45

50

55

EP 2 566 497 B1

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

<210> 51

<211> 759

<212> DNA

<213> Artificial Sequence

<220>

<223> Variant of L. kefir ketoreductase

<400> 51

EP 2 566 497 B1

	atgaccaacc gtctgaagag caaagtagcc atcgtaaccg gcgggaccca gggatatcgg	60
	ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttatcac cggtcgccgt	120
5	gcagatgtag gtgaaaaggc cgccaaatca atcggcggta ctgatgttat tcgctttgtc	180
	cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca	240
	ttcggcccggtt ttacgaccgt cgtgaacaat gcaggggcct ctctgaacaa aagcgttgaa	300
10	gacactacca cggaggaatg gcgcaaactg ttgtccgtta atctggatag tgtttttttc	360
	ggcaccgcgtc tgggcattcg ccgcatgaag aataaaggct tgggcgctag catcatcaat	420
15	atgagctcga tcagtgggct ggtaggcgat ccgatgttgg gggcatacaa cgcttccaag	480
	ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat	540
	gtgcgtgtca acacagtaca tccgggcgct atcaagacc cgctgactga taaatttcca	600
20	ggtggtgggg aaatgatctc acagcgtagc agaacccta tgggccacat tggcgaaccg	660
	aatgacgtgg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt	720
25	agcgaatttg tggtcgacgg cgggtatacc gcacagtga	759

<210> 52

<211> 252

<212> PRT

30 <213> Artificial Sequence

<220>

<223> Variant of L. kefir ketoreductase

35 <400> 52

40

45

50

55

EP 2 566 497 B1

	Met	Thr	Asn	Arg	Leu	Lys	Ser	Lys	Val	Ala	Ile	Val	Thr	Gly	Gly	Thr	
	1				5					10					15		
5	Gln	Gly	Ile	Gly	Leu	Ala	Ile	Ala	Asp	Lys	Phe	Val	Glu	Glu	Gly	Ala	
				20					25					30			
10	Lys	Val	Val	Ile	Thr	Gly	Arg	Arg	Ala	Asp	Val	Gly	Glu	Lys	Ala	Ala	
			35					40					45				
	Lys	Ser	Ile	Gly	Gly	Thr	Asp	Val	Ile	Arg	Phe	Val	Gln	His	Asp	Ala	
		50					55					60					
15	Ser	Asp	Glu	Ala	Gly	Trp	Thr	Lys	Leu	Phe	Asp	Thr	Thr	Glu	Glu	Ala	
	65					70					75					80	
20	Phe	Gly	Pro	Val	Thr	Thr	Val	Val	Asn	Asn	Ala	Gly	Ala	Ser	Leu	Asn	
					85					90					95		
	Lys	Ser	Val	Glu	Asp	Thr	Thr	Thr	Glu	Glu	Trp	Arg	Lys	Leu	Leu	Ser	
25				100					105					110			
	Val	Asn	Leu	Asp	Ser	Val	Phe	Phe	Gly	Thr	Arg	Leu	Gly	Ile	Arg	Arg	
			115					120					125				
30	Met	Lys	Asn	Lys	Gly	Leu	Gly	Ala	Ser	Ile	Ile	Asn	Met	Ser	Ser	Ile	
		130					135					140					
35	Ser	Gly	Leu	Val	Gly	Asp	Pro	Met	Leu	Gly	Ala	Tyr	Asn	Ala	Ser	Lys	
	145					150					155					160	
40	Gly	Ala	Val	Arg	Ile	Met	Ser	Lys	Ser	Ala	Ala	Leu	Asp	Cys	Ala	Val	

EP 2 566 497 B1

						165				170					175		
5		Lys	Asp	Tyr	Asp	Val	Arg	Val	Asn	Thr	Val	His	Pro	Gly	Ala	Ile	Lys
					180					185					190		
10		Thr	Pro	Leu	Thr	Asp	Lys	Phe	Pro	Gly	Gly	Gly	Glu	Met	Ile	Ser	Gln
				195					200					205			
15		Arg	Thr	Arg	Thr	Pro	Met	Gly	His	Ile	Gly	Glu	Pro	Asn	Asp	Val	Ala
			210					215					220				
20		Trp	Ile	Cys	Val	Tyr	Leu	Ala	Ser	Asp	Glu	Ser	Lys	Phe	Ala	Thr	Gly
		225					230					235					240
25		Ser	Glu	Phe	Val	Val	Asp	Gly	Gly	Tyr	Thr	Ala	Gln				
					245						250						
30		<210> 53															
		<211> 759															
		<212> DNA															
35		<213> Artificial Sequence															
		<220>															
		<223> Variant of L. kefir ketoreductase															
40		<400> 53															
45		atgaccaacc	gtctgaagag	caaagtagcc	atcgtaaccg	gcgggaccca	gggtatcggc										60
50		ttggcaatcg	ccactaaatt	tgtagaggag	ggtgcgaaag	tagttatcac	cggtcgccgt										120
55		gcagatgtag	gtgaaaaggc	cgccaaatca	atcggcggta	ctgatgttat	tcgctttgtc										180
60		cagcacgatg	catccgatga	agcaggctgg	acgaaactgt	tcgacaccac	cgaggaggca										240
65		ttcggcccgg	ttacgaccgt	cgtgaacaat	gcagggggcca	ctctgcagaa	aagcgttgaa										300
70		gacactacca	cggaggaatg	gcacaaactg	ttgtccgtta	atctggatag	tgtttttttc										360
75		ggcaccgcgc	tgggcattcg	ccgcatgaag	aataaaggct	tgggcgctag	catcatcaat										420
80		atgagctcga	tcagtgggct	ggtaggcgat	ccgatgatcg	gggcatacaa	tgcttccaag										480
85		ggggcggtac	gtatcatgtc	gaaaagcgca	gcgctggatt	gcgcagtgaa	ggactacgat										540
90		gtgcgtgtca	acacagtaca	tccggggcgt	atcaagaccc	cgctgactga	taaatttcca										600
95		ggtgctgggg	aatgatctc	acagcgtacg	agaacccta	tgggccacat	tggcgaaccg										660
100		aatgacgtgg	catggatctg	tgtgtacctg	gcacatctgacg	aatcgaaatt	tgcgacgggt										720
105		agcgaatttg	tggtcgacgg	cggtatatacc	gcacagtga												759
110		<210> 54															
		<211> 252															

EP 2 566 497 B1

<212> PRT

<213> Artificial Sequence

<220>

5 <223> Variant of L. kefir ketoreductase

<400> 54

10

15

20

25

30

35

40

45

50

55

EP 2 566 497 B1

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

245

250

5 <210> 55
 <211> 759
 <212> DNA
 <213> Artificial Sequence

10 <220>
 <223> Variant of L. kefir ketoreductase

<400> 55

15 atgaccaacc gtctgaagag caaagtagcc atcgtaaccg gcgggaccca gggatatcgg 60
 ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttatcac cggtcgccgt 120
 gcagatgtag gtgaaaaggc cgccaaatca atcggcggtta ctgatgttat tcgctttgtc 180
 20 cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
 ttccggcccg ttacgaccgt cgtgaacaat gcaggggcca ctctgcagaa aagcgttgaa 300
 gacactacca cggaggaatg gcgcaaactg ttgtccgtta atctggatag tgtttttttc 360
 25 ggcaccgctc tgggcattcg ccgcatgaag aataaaggct tgggcgctag catcatcaat 420
 atgagctcga tcagtgggct ggtaggccat ccgatgatcg gggcatacaa tgcttccaag 480
 30 ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat 540
 gtgcgtgtca acacagtaca tccgggcgct atcaagacc cgctgactga taaatttcca 600
 ggtgctgggg aaatgatctc acagcgtacg agaacccta tgggccacat tggcgaaccg 660
 35 aatgacgtgg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
 agcgaatttg tggtcgacgg cgggtatacc gcacagtga 759

40 <210> 56
 <211> 252
 <212> PRT
 <213> Artificial Sequence

45 <220>
 <223> Variant of L. kefir ketoreductase

<400> 56

50

55

EP 2 566 497 B1

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

EP 2 566 497 B1

<210> 57
 <211> 759
 <212> DNA
 <213> Artificial Sequence

5

<220>
 <223> Variant of L. kefir ketoreductase

<400> 57

10

atgaccaacc	gtctgaagag	caaagtagcc	atcgtaaccg	gcgggaccca	gggtatcgg	60
ttggcaatcg	ccgataaatt	tgtagaggag	ggtgcgaaag	tagttatcac	cggtcgcgct	120
gcagatgtag	gtgaaaaggc	cgccaaatca	atcggcggtta	ctgatgttat	tcgctttgtc	180
cagcacgatg	catccgatga	agcaggctgg	acgaaactgt	tcgacaccac	cgaggaggca	240
ttcggcccg	ttacgaccgt	cgtgaacaat	gcaggggcca	ctctgcagaa	aagcgttgaa	300
gacactacca	cggaggaatg	gcgcaaactg	ttgtccgtta	atctggatag	tgtttttttc	360
ggcaccgctc	tgggcattcg	ccgcatgaag	aataaaggct	tgggcgctag	catcatcaat	420
atgagctcga	tcagtgggct	ggtaggcgat	ccgatgatcg	gggcatacaa	tgcttccaag	480
ggggcggtac	gtatcatgtc	gaaaagcgca	gcgctggatt	gcgcagtga	ggactacgat	540
gtgcgtgtca	acacagtaca	tccggggccg	atcaagaccc	cgctgactga	taaatttcca	600
ggtgctgggg	aaatgatctc	acagcgtagc	agaaccctta	tgggccacat	tggcgaaccg	660
aatgacgtgg	catggatctg	tgtgtacctg	gcattctgacg	aatcgaaatt	tgcgacgggt	720
agcgaatttg	tggtcgacgg	cgggtatacc	gcacagtga			759

35

<210> 58
 <211> 252
 <212> PRT
 <213> Artificial Sequence

40

<220>
 <223> Variant of L. kefir ketoreductase

45

<400> 58

50

55

EP 2 566 497 B1

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

EP 2 566 497 B1

<210> 59
 <211> 759
 <212> DNA
 <213> Artificial Sequence

5

<220>
 <223> Variant of L. kefir ketoreductase

<400> 59

10

atgaccaacc	gtctgaagag	caaagtagcc	atcgtaaccg	gcgggaccca	gggtatcggt	60
ttggcaatcg	ccgataaatt	tgtagaggag	ggtgcgaaag	tagttatcac	cggtcgccgt	120
gcagatgtag	gtgaaaaggc	cgccaaatca	atcggcggtg	ctgatgttat	tcgctttgtc	180
cagcacgatg	catccgatga	agcaggctgg	acgaaactgt	tcgacaccac	cgaggaggca	240
ttcggcccgg	ttacgaccgt	cgtgaacaat	gcaggggcca	ctctgcagaa	aagcgttgaa	300
gacactacca	cggaggaatg	gcgcaaactg	ttgtccgtta	atctggatag	tgtttttttc	360
ggcaccgcgc	tgggcattcg	ccgcatgaag	aataaaggct	tgggcgctag	catcatcaat	420
atgagctcga	tcagtgggct	ggtaggcgat	ccgatgatcg	gggcatacaa	tgcttccaag	480
ggggcggtac	gtatcatgtc	gaaaagcgca	gcgctggatt	gcgcagtga	ggactacgat	540
gtgcgtgtca	acacagtaca	tccgggcgct	atcaagacc	cgctgactga	taaatttcca	600
ggtgctgggg	caatgatctc	acagcgtag	agaacccta	tgggccacat	tggcgaaccg	660
aatgacgtgg	catggatctg	tgtgtacctg	gcctctgacg	aatcgaaatt	tgcgacgggt	720
agcgaatttg	tggtcgacgg	cgggtatacc	gcacagtga			759

35

<210> 60
 <211> 252
 <212> PRT
 <213> Artificial Sequence

40

<220>
 <223> Variant of L. kefir ketoreductase

<400> 60

45

50

55

EP 2 566 497 B1

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

EP 2 566 497 B1

<210> 61
 <211> 759
 <212> DNA
 <213> Artificial Sequence

5

<220>
 <223> Variant of L. kefir ketoreductase

<400> 61

10

```

atgaccaacc gtctgaagag caaagtagcc atcgtaaccg gcgggaccca gggatatcggt      60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttatcac cggtcgccgt      120
gcagatgtag gtgaaaaggc cgccaaatca atcggcggta ctgatgttat tcgctttgtc      180
cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca      240
ttcggccccg ttacgaccgt cgtgaacaat gcagggggcca ctctgcagaa aagcgttgaa      300
gacactacca cggaggaatg gcgcaaactg ttgtccgtta atctggatag tgtttttttc      360
ggcaccgcgc tgggcattcg ccgcatgaag aataaaggct tgggcgctag catcatcaat      420
atgagctcga tcagtgggct ggtaggcgat ccgatgatcg gggcatacaa tgcttccaag      480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat      540
gtgcgtgtca acacagtaca tccgggcgct atcaagaccg cgctgactga taaatttcca      600
gggtgctggtg taatgatctc acagcgtacg agaacccta tgggccacat tggcgaaccg      660
aatgacgtgg catggatctg tgtgtacctg gcctctgacg aatcgaaatt tgcgacgggt      720
agcgaatttg tggtcgacgg cgggtatacc gcacagtga      759
    
```

35

<210> 62
 <211> 252
 <212> PRT
 <213> Artificial Sequence

40

<220>
 <223> Variant of L. kefir ketoreductase

<400> 62

45

```

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

Gln Gly Ile Gly Leu Ala Ile Ala Asp Lys Phe Val Glu Glu Gly Ala
                20              25              30

Lys Val Val Ile Thr Gly Arg Arg Ala Asp Val Gly Glu Lys Ala Ala
35              40              45
    
```

55

EP 2 566 497 B1

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

<210> 63

50 <211> 759

<212> DNA

<213> Artificial Sequence

<220>

55 <223> Variant of L. kefir ketoreductase

<400> 63

atgaccaacc gtctgaagag caaagtagcc atcgtaaccg gcgggaccca gggatatcgg 60
 ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttatcac cggtcgccgt 120
 5 gcagatgtag gtgaaaaggc cgccaaatca atcggcggtta ctgatgttat tcgctttgtc 180
 cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
 10 ttcgggcccg ttacgaccgt cgtgaacaat gcaggggcca ctctgcagaa aagcgttgaa 300
 gacactacca cggaggaatg gcgcaaactg ttgtccgtta atctggatag tgtttttttc 360
 ggcacccgtc tgggcattcg ccgcatgaag aataaaggct tgggcgctag catcatcaat 420
 15 atgagctcga tcagtgggct ggtaggcgat ccgatgatcg gggcatacaa tgcttccaag 480
 ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat 540
 20 gtgcgtgtca acacagtaca tccgggcgct atcaagacct cgctgactga taaatttcca 600
 ggtgctgggg aagtgatctc acagcgtacg agaacccta tgggccacat tggcgaaccg 660
 aatgacgtgg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
 25 agcgaatttg tggtcgacgg cgggtatacc gcacagtga 759

<210> 64

<211> 252

<212> PRT

30 <213> Artificial Sequence

<220>

<223> Variant of L. kefir ketoreductase

35 <400> 64

40

45

50

55

EP 2 566 497 B1

	Met	Thr	Asn	Arg	Leu	Lys	Ser	Lys	Val	Ala	Ile	Val	Thr	Gly	Gly	Thr	
	1				5					10					15		
5	Gln	Gly	Ile	Gly	Leu	Ala	Ile	Ala	Asp	Lys	Phe	Val	Glu	Glu	Gly	Ala	
				20					25					30			
10	Lys	Val	Val	Ile	Thr	Gly	Arg	Arg	Ala	Asp	Val	Gly	Glu	Lys	Ala	Ala	
			35					40					45				
15	Lys	Ser	Ile	Gly	Gly	Thr	Asp	Val	Ile	Arg	Phe	Val	Gln	His	Asp	Ala	
		50					55					60					
20	Ser	Asp	Glu	Ala	Gly	Trp	Thr	Lys	Leu	Phe	Asp	Thr	Thr	Glu	Glu	Ala	
	65					70					75					80	
25	Phe	Gly	Pro	Val	Thr	Thr	Val	Val	Asn	Asn	Ala	Gly	Ala	Thr	Leu	Gln	
					85					90					95		
30	Lys	Ser	Val	Glu	Asp	Thr	Thr	Thr	Glu	Glu	Trp	Arg	Lys	Leu	Leu	Ser	
				100					105					110			
35	Val	Asn	Leu	Asp	Ser	Val	Phe	Phe	Gly	Thr	Arg	Leu	Gly	Ile	Arg	Arg	
			115					120					125				

EP 2 566 497 B1

	Met	Lys	Asn	Lys	Gly	Leu	Gly	Ala	Ser	Ile	Ile	Asn	Met	Ser	Ser	Ile
	130						135					140				
5	Ser	Gly	Leu	Val	Gly	Asp	Pro	Met	Ile	Gly	Ala	Tyr	Asn	Ala	Ser	Lys
	145					150					155					160
10	Gly	Ala	Val	Arg	Ile	Met	Ser	Lys	Ser	Ala	Ala	Leu	Asp	Cys	Ala	Val
					165					170					175	
15	Lys	Asp	Tyr	Asp	Val	Arg	Val	Asn	Thr	Val	His	Pro	Gly	Ala	Ile	Lys
				180					185					190		
20	Thr	Pro	Leu	Thr	Asp	Lys	Phe	Pro	Gly	Ala	Gly	Glu	Val	Ile	Ser	Gln
			195					200					205			
25	Arg	Thr	Arg	Thr	Pro	Met	Gly	His	Ile	Gly	Glu	Pro	Asn	Asp	Val	Ala
	210						215					220				
30	Trp	Ile	Cys	Val	Tyr	Leu	Ala	Ser	Asp	Glu	Ser	Lys	Phe	Ala	Thr	Gly
	225					230					235					240
35	Ser	Glu	Phe	Val	Val	Asp	Gly	Gly	Tyr	Thr	Ala	Gln				
					245					250						

<210> 65

<211> 759

<212> DNA

<213> Artificial Sequence

<220>

<223> Variant of L. kefir ketoreductase

<400> 65

	atgaccaacc gtctgaagag caaagtagcc atcgtaaccg gcgggaccca gggatatcgg	60
	ttggcaatcg ccgataaatt thtagaggag ggtgcgaaag tagttatcac cggtcgccgt	120
5	gcagatgtag gtgaaaaggc cgccaaatca atcggcggtta ctgatgttat tcgctttgtc	180
	cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca	240
	ttcggccccg ttacgaccgt cgtgaacaat gcaggggcca ctctgcagaa aagcgttgaa	300
10	gacactacca cggaggaatg gcgcaaactg ttgtccgtta atctggatag tgtttttttc	360
	ggcaccgcgtc tgggcattcg ccgcatgaag aataaaggct tgggcgctag catcatcaat	420
	atgagctcga tcagtgggct ggtaggcgat ccgaatatcg gggcatacaa tgcttccaag	480
15	ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat	540
	gtgcgtgtca acacagtaca tccgggcgct atcaagacc cgctgactga taaatttcca	600
20	ggtgctgggg aaatgatctc acagcgtacg agaacccta tgggccacat tggcgaaccg	660
	aatgacgtgg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt	720
25	agcgaatttg tggtcgacgg cgggtatacc gcacagtga	759

<210> 66

<211> 252

<212> PRT

30 <213> Artificial Sequence

<220>

<223> Variant of L. kefir ketoreductase

35 <400> 66

40

45

50

55

EP 2 566 497 B1

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

EP 2 566 497 B1

<210> 67
 <211> 759
 <212> DNA
 <213> Artificial Sequence

5

<220>
 <223> Variant of L. kefir ketoreductase

<400> 67

10

```

atgaccaacc gtctgaagag caaagtagcc atcgtaaccc gcgggaccca gggatatcgg      60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttatcac cggtcgccgt      120
gcagatgtag gtgaaaaggc cgccaaatca atcggcggta ctgatgttat tcgctttgtc      180
cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca      240
ttcggccccg ttacgaccgt cgtgaacaat gcagggggcca ctctgcagaa aagcgttgaa      300
gacactacca cggaggaatg gcgcaaactg ttgtccgta atctggatag tgtttttttc      360
ggcaccgcgc tgggcattcg ccgcatgaag aataaaggct tgggcgctag catcatcaat      420
atgagctcga tcagtgggct ggtaggcgat ccgtttatcg gggcatacaa tgcttccaag      480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat      540
gtgcgtgtca acacagtaca tccgggcgct atcaagacc cgctgactga taaatttcca      600
gggtgctggg aaatgatctc acagcgtacg agaacccta tgggccacat tggcgaaccg      660
aatgacgtgg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt      720
agcgaatttg tggtcgacgg cgggtatacc gcacagtga                          759
    
```

35

<210> 68
 <211> 252
 <212> PRT
 <213> Artificial Sequence

40

<220>
 <223> Variant of L. kefir ketoreductase

<400> 68

45

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

50

Gln Gly Ile Gly Leu Ala Ile Ala Asp Lys Phe Val Glu Glu Gly Ala
 20 25 30

55

EP 2 566 497 B1

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

<210> 69

<211> 759

55 <212> DNA

<213> Artificial Sequence

<220>

EP 2 566 497 B1

<223> Variant of L. kefir ketoreductase

<400> 69

```

5  atgaccaacc gtctgaagag caaagtagcc atcgtaaccg gcgggaccca gggatatcgg 60
   ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttatcac cggtcgccgt 120
   gcagatgtag gtgaaaaggc cgccaaatca atcggcggta ctgatgttat tcgctttgtc 180
10 cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
   ttcgggcccg ttacgaccgt cgtgaacaat gcagggggcca ctctggcaaa aagcgttgaa 300
15 gacactacca cggaggaatg gcgcaaactg ttgtccgtta atctggatag tgtttttttc 360
   ggcacccgtc tgggcattcg ccgcatgaag aataaaggct tgggcgctag catcatcaat 420
   atgagctcga tcagtgggct ggtaggcgat ccgatgatcg gggcatacaa tgcttccaag 480
20 ggggcgggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat 540
   gtgcgtgtca acacagtaca tccgggcgct atcaagacc cgctgactga taaatttcca 600
   ggtgctgggg aaatgatcac ccagcgtacg agaacccta tgggccacat tggcgaaccg 660
25 aatgacgtgg catggatctg tgtgtacctg gcctctgacg aatcgaaatt tgcgacgggt 720
   agcgaatttg tggtcgacgg cgggtatacc gcacagtga 759

```

```

30  <210> 70
    <211> 252
    <212> PRT
    <213> Artificial Sequence

```

```

35  <220>
    <223> Variant of L. kefir ketoreductase

    <400> 70

```

40

45

50

55

EP 2 566 497 B1

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

EP 2 566 497 B1

<210> 71
 <211> 759
 <212> DNA
 <213> Artificial Sequence

5

<220>
 <223> Variant of L. kefir ketoreductase

<400> 71

10

atgaccaacc	gtctgaagag	caaagtagcc	atcgtaaccg	gcgggaccca	gggtatcggg	60
ttggcaatcg	ccaacaaatt	tgtagaggag	ggtgcgaaag	tagttatcac	cggtcgccgt	120
gcagatgtag	gtgaaaaggc	cgccaaatca	atcggcggta	cggatgttat	tcgctttgtc	180
cagcacgatg	catccgatga	agcaggctgg	acgaaactgt	tcgacaccac	cgaggaggca	240
ttcggccccg	ttacgaccgt	cgtgaacaat	gcagggggcca	ctctgcagaa	aagcgttgaa	300
gacactacca	cggaggaatg	gcgcaaactg	ttgtccgtta	atctggatag	tgtttttttc	360
ggcaccgcgc	tgggcattcg	ccgcatgaag	aataaaggct	tgggcgctag	catcatcaat	420
atgagctcga	tcagtgggct	ggtaggcgat	ccgatgatcg	gggcatacaa	tgcttccaag	480
ggggcggtac	gtatcatgtc	gaaaagcgca	gcgctggatt	gcgcagtgaa	ggactacgat	540
gtgcgtgtca	acacagtaca	tccgggcgct	atcaagaccc	cgctgactga	taaatttcca	600
ggtgctgggg	aaatgatctc	acagcgtacg	agaacccta	tgggccacat	tggcgaaccg	660
aatgacgtgg	catggatctg	tgtgtacctg	gcattctgacg	aatcgaaatt	tgcgacgggt	720
agcgaatttg	tggtcgacgg	cggtatatacc	gcacagtga			759

35

<210> 72
 <211> 252
 <212> PRT
 <213> Artificial Sequence

40

<220>
 <223> Variant of L. kefir ketoreductase

<400> 72

45

50

55

EP 2 566 497 B1

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

EP 2 566 497 B1

<210> 73
 <211> 759
 <212> DNA
 <213> Artificial Sequence

5

<220>
 <223> Variant of L. kefir ketoreductase

<400> 73

10

atgaccaacc	gtctgaagag	caaagtagcc	atcgtaaccg	gcgggaccca	gggtatcggt	60
ttggcaatcg	ccacaaaatt	tgtagaggag	ggtgcgaaag	tagttatcac	cggtcgccgt	120
gcagatgtag	gtgaaaaggc	cgccaaatca	atcggcggtg	ctgatgttat	tcgctttgtc	180
cagcacgatg	catccgatga	agcaggctgg	acgaaactgt	tcgacaccac	cgaggaggca	240
ttcggcccgg	ttacgaccgt	cgtgaacaat	gcagggggcca	ctctgcagaa	aagcgttgaa	300
gacactacca	cggaggaatg	gcgcaaactg	ttgtccgtta	atctggatag	tgtttttttc	360
ggcaccgcgc	tgggcattcg	ccgcatgaag	aataaaggct	tgggcgctag	catcatcaat	420
atgagctcga	tcagtgggct	ggtaggcgat	ccgatgatcg	gggcatacaa	tgcttccaag	480
ggggcggtac	gtatcatgtc	gaaaagcgca	gcgctggatt	gcgcagtga	ggactacgat	540
gtgcgtgtca	acacagtaca	tccgggcgct	atcaagaccc	cgctgactga	taaatttcca	600
ggtgctgggg	aaatgatctc	acagcgtacg	agaacccta	tgggccacat	tggcgaaccg	660
aatgacgtgg	catggatctg	tgtgtacctg	gcacctgacg	aatcgaaatt	tgcgacgggt	720
agcgaatttg	tggtcgacgg	cgggtatacc	gcacagtga			759

35

<210> 74
 <211> 252
 <212> PRT
 <213> Artificial Sequence

40

<220>
 <223> Variant of L. kefir ketoreductase

<400> 74

45

Met Thr Asn Arg Leu Lys Ser Lys Val Ala Ile Val Thr Gly Gly Thr

50

55

EP 2 566 497 B1

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

EP 2 566 497 B1

<210> 75
 <211> 759
 <212> DNA
 <213> Artificial Sequence

5

<220>
 <223> Variant of L. kefir ketoreductase

<400> 75

10

atgaccaacc	gtctgaagag	caaagtagcc	atcgtaaccg	gcgggaccca	gggtatcggt	60
ttggcaatcg	ccgataaatt	tgtagaggag	ggtgcgaaag	tagttatcac	cggtcgccgt	120
gcagatgtag	gtgaaaaggc	cgccaaatca	atcggcggta	ctgatgttat	tcgctttgtc	180
cagcacgatg	catccgatga	agcaggctgg	acgaaactgt	tcgacaccac	cgaggaggca	240
ttcggcccgg	ttacgaccgt	cgtgaacaat	gcagggggcca	ctctgcagaa	aagcgttgaa	300
gacactacca	cggaggaatg	gcgcaaactg	ttgtccgtta	atctggatag	tgtttttttc	360
ggcaccgcgc	tgggcattcg	ccgcatgaag	aataaaggct	tgggcgctag	catcatcaat	420
atgagctcga	tcagtgggct	ggtaggcgat	ccgatgatcg	gggcatacaa	tgcttccaag	480
ggggcggtac	gtatcatgtc	gaaaagcgca	gcgctggatt	gcgcagtga	ggactacgat	540
gtgcgtgtca	acacagtaca	tccgggcgct	atcaagacc	cgatgactga	taaatttcca	600
ggtgctgggg	aaatgatctg	ccagcgtacg	agaacccta	tgggccacat	tggcgaaccg	660
aatgacgtgg	catggatctg	tgtgtacctg	gcactctgacg	aatcgaaatt	tgcgacgggt	720
agcgaatttg	tggtcgacgg	cgggtatacc	gcacagtga			759

35

<210> 76
 <211> 252
 <212> PRT
 <213> Artificial Sequence

40

<220>
 <223> Variant of L. kefir ketoreductase

<400> 76

45

50

55

EP 2 566 497 B1

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

EP 2 566 497 B1

85

90

95

5 Lys Ser Val Glu Asp Thr Thr Thr Glu Glu Trp Arg Lys Leu Leu Ser
100 105 110

10 Val Asn Leu Asp Ser Val Phe Phe Gly Thr Arg Leu Gly Ile Arg Arg
115 120 125

Met Lys Asn Lys Gly Leu Gly Ala Ser Ile Ile Asn Met Ser Ser Ile
130 135 140

15 Ser Gly Leu Val Gly Asp Pro Met Ile Gly Ala Tyr Asn Ala Ser Lys
145 150 155 160

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

25 Lys Asp Tyr Asp Val Arg Val Asn Thr Val His Pro Gly Ala Ile Lys
180 185 190

Thr Pro Met Thr Asp Lys Phe Pro Gly Ala Gly Glu Met Ile Cys Gln
195 200 205

30 Arg Thr Arg Thr Pro Met Gly His Ile Gly Glu Pro Asn Asp Val Ala
210 215 220

35 Trp Ile Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly
225 230 235 240

40 Ser Glu Phe Val Val Asp Gly Gly Tyr Thr Ala Gln
245 250

<210> 77

<211> 759

<212> DNA

<213> Artificial Sequence

<220>

<223> Variant of L. kefir ketoreductase

<400> 77

EP 2 566 497 B1

	atgaccaacc gtctgaagag caaagtagcc atcgtaaccg gcgggaccca gggatcgggt	60
	ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttatcac cggtcgccgt	120
5	gcagatgtag gtgaaaaggc cgccaaatca atcggcggta ctgatgttat tcgctttgtc	180
	cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca	240
	ttcggcccgg ttacgaccgt cgtgaacaat gcaggggcct ctctgagcaa aagcgttgaa	300
10	gacactacca cggaggaatg gcgcaaactg ttgtccgtta atctggatag tgtttttttc	360
	ggcaccgcgc tgggcattcg ccgcatgaag aataaaggct tgggcgctag catcatcaat	420
15	atgagctcga tcagtgggct ggtaggcgat ccgatgttgg gggcatacaa cgcttccaag	480
	ggggcgggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat	540
	gtgcgtgtca acacagtaca tccgggcgct atcaagaccc cgctgactga taaatttcca	600
20	ggtggtgggg aaatgatctc acagcgtacg cgcacccta tgggccacat tggatgaaccg	660
	aatgacgtgg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt	720
25	agcgaatttg tggtcgacgg cgggtatacc gcacagtga	759

<210> 78

<211> 252

<212> PRT

30 <213> Artificial Sequence

<220>

<223> Variant of L. kefir ketoreductase

35 <400> 78

40

45

50

55

EP 2 566 497 B1

	Met	Thr	Asn	Arg	Leu	Lys	Ser	Lys	Val	Ala	Ile	Val	Thr	Gly	Gly	Thr	
	1				5					10					15		
5	Gln	Gly	Ile	Gly	Leu	Ala	Ile	Ala	Asp	Lys	Phe	Val	Glu	Glu	Gly	Ala	
				20					25					30			
10	Lys	Val	Val	Ile	Thr	Gly	Arg	Arg	Ala	Asp	Val	Gly	Glu	Lys	Ala	Ala	
			35					40					45				
	Lys	Ser	Ile	Gly	Gly	Thr	Asp	Val	Ile	Arg	Phe	Val	Gln	His	Asp	Ala	
		50					55					60					
15	Ser	Asp	Glu	Ala	Gly	Trp	Thr	Lys	Leu	Phe	Asp	Thr	Thr	Glu	Glu	Ala	
	65					70					75					80	
20	Phe	Gly	Pro	Val	Thr	Thr	Val	Val	Asn	Asn	Ala	Gly	Ala	Ser	Leu	Ser	
					85					90					95		
	Lys	Ser	Val	Glu	Asp	Thr	Thr	Thr	Glu	Glu	Trp	Arg	Lys	Leu	Leu	Ser	
25				100					105					110			
	Val	Asn	Leu	Asp	Ser	Val	Phe	Phe	Gly	Thr	Arg	Leu	Gly	Ile	Arg	Arg	
			115					120					125				
30	Met	Lys	Asn	Lys	Gly	Leu	Gly	Ala	Ser	Ile	Ile	Asn	Met	Ser	Ser	Ile	
		130					135					140					
35	Ser	Gly	Leu	Val	Gly	Asp	Pro	Met	Leu	Gly	Ala	Tyr	Asn	Ala	Ser	Lys	
	145					150					155					160	
40	Gly	Ala	Val	Arg	Ile	Met	Ser	Lys	Ser	Ala	Ala	Leu	Asp	Cys	Ala	Val	

EP 2 566 497 B1

						165				170					175	
5	Lys	Asp	Tyr	Asp	Val	Arg	Val	Asn	Thr	Val	His	Pro	Gly	Ala	Ile	Lys
				180					185					190		
10	Thr	Pro	Leu	Thr	Asp	Lys	Phe	Pro	Gly	Gly	Gly	Glu	Met	Ile	Ser	Gln
			195					200					205			
15	Arg	Thr	Arg	Thr	Pro	Met	Gly	His	Ile	Gly	Glu	Pro	Asn	Asp	Val	Ala
	210						215					220				
20	Trp	Ile	Cys	Val	Tyr	Leu	Ala	Ser	Asp	Glu	Ser	Lys	Phe	Ala	Thr	Gly
	225					230					235					240
25	Ser	Glu	Phe	Val	Val	Asp	Gly	Gly	Tyr	Thr	Ala	Gln				
				245						250						
30	<210> 79															
	<211> 759															
	<212> DNA															
35	<213> Artificial Sequence															
	<220>															
	<223> Variant of L. kefir ketoreductase															
40	<400> 79															
45	atgaccaacc	gtctgaagag	caaagtagcc	atcgtaaccg	gcgggaccca	gggtatcggt										60
50	ttggcaatcg	ccgataaatt	tgtagaggag	gggtgcgaaag	tagttatcac	cggtcgccgt										120
55	gcagatgtag	gtgaaaaggc	cgccaaatca	atcggcggtg	ctgatgttat	tcgctttgtc										180
60	cagcacgatg	catccgatga	agcaggctgg	acgaaactgt	tcgacaccac	cgaggaggca										240
65	ttcggcccgg	ttacgaccgt	cgtgaacaat	gcaggggcct	ctctgcctaa	aagcgttgaa										300
70	gacactacca	cggaggaatg	gcgcaaactg	ttgtccgtta	atctggatag	tgtttttttc										360
75	ggcaccgcgc	tgggcattcg	ccgcatgaag	aataaaggct	tgggcgctag	catcatcaat										420
80	atgagctcga	tcagtgggct	ggtaggcgat	ccgatgttgg	gggcatacaa	cgcttccaag										480
85	ggggcggtac	gtatcatgtc	gaaaagcgca	gcgctggatt	gcgcagtgaa	ggactacgat										540
90	gtgctgtca	acacagtaca	tccgggcgct	atcaagaccc	cgctgactga	taaatttcca										600
95	ggcgggtggg	aatgatctc	acagcgtagc	agaacccta	tgggccacat	tggtgaaccg										660
100	aatgacgtgg	catggatctg	tgtgtacctg	gcacgtgacg	aatcgaaatt	tgcgacgggt										720
105	agcgaatttg	tggtcgacgg	cgggtatacc	gcacagtga												759
110	<210> 80															
115	<211> 252															

EP 2 566 497 B1

<212> PRT

<213> Artificial Sequence

<220>

5 <223> Variant of L. kefir ketoreductase

<400> 80

10

15

20

25

30

35

40

45

50

55

EP 2 566 497 B1

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

245

250

5 <210> 81
 <211> 759
 <212> DNA
 <213> Artificial Sequence

 10 <220>
 <223> Variant of L. kefir ketoreductase

 <400> 81

 15 atgaccaacc gtctgaagag caaagtagcc atcgtaaccg gcgggaccca gggatatcgg 60
 ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttatcac cggtcgccgt 120
 gcagatgtag gtgaaaaggc cgccaaatca atcggcggta ctgatgttat tcgctttgtc 180
 20 cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
 ttccggcccg ttacgaccgt cgtgaacaat gcaggggcct ctctggcgaa aagcgttgaa 300
 gacactacca cggaggaatg gcgcaaactg ttgtccgta atctggatag tgtttttttc 360
 25 ggcaccogtc tgggcattcg ccgcatgaag aataaaggct tgggcgctag catcatcaat 420
 atgagctcga tcagtgggct ggtaggcgat ccgatgttg gggcatacaa cgcttccaag 480
 30 ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat 540
 gtgcgtgtca acacagtaca tccgggcgct atcaagacc cgctgactga taaatttcca 600
 ggtggtgggg aaatgatctc acagcgtacg agaacccta tgggccacat tggngaaccg 660
 35 aatgacgtgg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
 agcgaatttg tggtcgacgg cgggtatacc gcacagtga 759

40 <210> 82
 <211> 252
 <212> PRT
 <213> Artificial Sequence

45 <220>
 <223> Variant of L. kefir ketoreductase

 <400> 82

50

55

EP 2 566 497 B1

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

EP 2 566 497 B1

<210> 83
<211> 759
<212> DNA
<213> Artificial Sequence

5

<220>
<223> Variant of L. kefir ketoreductase

<400> 83

10

atgaccaacc	gtctgaagag	caaagtagcc	atcgtaaccg	gcgggaccca	gggtatcgg	60
ttggcaatcg	ccgataaatt	tgtagaggag	ggtgcgaaag	tagttatcac	cggtcgccgt	120
gcagatgtag	gtgaaaaggc	cgccaaatca	atcggcggt	ctgatgttat	tcgctttgtc	180
cagcacgatg	catccgatga	agcaggctgg	acgaaactgt	tcgacaccac	cgaggaggca	240
ttcggcccg	ttacgaccgt	cgtgaacaat	gcaggggcct	ctctgacgaa	aagcgttgaa	300
gacactacca	cggaggaatg	gcgcaaactg	ttgtccgtta	atctggatag	tgtttttttc	360
ggcaccctgc	tgggcattcg	ccgcatgaag	aataaaggct	tgggcgctag	catcatcaat	420
atgagctcga	tcagtgggct	ggtaggcgat	ccgatgttgg	gggcatacaa	cgtttccaag	480
ggggcggtac	gtatcatgtc	gaaaagcgca	gcgctggatt	gcgcagtga	ggactacgat	540
gtgcgtgtca	acacagtaca	tccgggcgct	atcaagacct	cgtgactga	taaattttcca	600
ggtggtgggg	aaatgatctc	acagcgtacg	agaaccctta	tgggccacat	tgggtgaaccg	660
aatgacgtgg	catggatctg	tgtgtacctg	gcctctgacg	aatcgaaatt	tgcgacgggt	720
agcgaatttg	tggtcgacgg	cgggtatacc	gcacagtga			759

35

<210> 84
<211> 252
<212> PRT
<213> Artificial Sequence

40

<220>
<223> Variant of L. kefir ketoreductase

<400> 84

45

50

55

EP 2 566 497 B1

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

EP 2 566 497 B1

<210> 85
 <211> 759
 <212> DNA
 <213> Artificial Sequence

5

<220>
 <223> Variant of L. kefir ketoreductase

<400> 85

10

```

atgaccaacc gtctgaagag caaagtagcc atcgtaaccg gcgggaccca gggatatcgg      60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttatcac cggtcgccgt      120
gcagatgtag gtgaaaaggc cgccaaatca atcggcggta ctgatgttat tcgctttgtc      180
cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca      240
ttcggcccgg ttacgaccgt cgtgaacaat gcaggggcct ctctgaacaa aagcgttgaa      300
gacactacca cggaggaatg gcgcaaactg ttgtccgtta atctggatag tgtttttttc      360
ggcaccgcgc tgggcattcg ccgcatgaag aataaaggct tgggcgctag catcatcaat      420
atgagctcga tcagtgggct ggtaggcgat ccgatgttgg gggcatacaa cgcttccaag      480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat      540
gtgcgtgtca acacagtaca tccgggcgct atcaagacc cgctgactga taaatttcca      600
gggtggtggg aaatgatctc acagcgtacg aagacccta tgggccacat tggatgaacc      660
aatgacgtgg catggatctg tgtgtacctg gcacgtgacg aatcgaaatt tgcgacgggt      720
agcgaatttg tggtcgacgg cgggtatacc gcacagtga      759
  
```

35

<210> 86
 <211> 252
 <212> PRT
 <213> Artificial Sequence

40

<220>
 <223> Variant of L. kefir ketoreductase

<400> 86

45

50

55

EP 2 566 497 B1

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

Trp Ile Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly
 225 230 235 240

5

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

<210> 87

10 <211> 759

<212> DNA

<213> Artificial Sequence

<220>

15 <223> Variant of L. kefir ketoreductase

<400> 87

20 atgaccaacc gtctgaagag caaagtagcc atcgtaaccg gcgggaccca gggatatcgggt 60
 ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttatcac cggtcgccgt 120
 gcagatgtag gtgaaaaggc cgccaaatca atcggcggtta ctgatgttat tcgctttgtc 180
 25 cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
 ttcgggcccggt ttacgaccgt cgtgaacaat gcaggggcct ctctgaacaa aagcgttgaa 300
 gacactacca cggaggaatg gcgcaaactg ttgtccgtta atctggatag tgtttttttc 360
 30 ggcacccgtc tgggcattcg ccgcatgaag aataaaggct tgggcgctag catcatcaat 420
 atgagctcga tcagtgggct ggtaggcgat ccgatgttgg ggtgttacia cgcttccaag 480
 35 ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat 540
 gtgcgtgtca acacagtaca tccgggcgct atcaagacct cgctgactga taaattttcca 600
 ggtggtgggg aaatgatctc acagcgtacg agaaccctta tgggccacat tggatgaaccg 660
 40 aatgacgtgg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
 agcgaatttg tggtcgacgg cgggtatacc gcacagtga 759

<210> 88

45 <211> 252

<212> PRT

<213> Artificial Sequence

<220>

50 <223> Variant of L. kefir ketoreductase

<400> 88

55

EP 2 566 497 B1

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

EP 2 566 497 B1

<210> 89
 <211> 759
 <212> DNA
 <213> Artificial Sequence

5

<220>
 <223> Variant of L. kefir ketoreductase

<400> 89

10

atgaccaacc	gtctgaagag	caaagtagcc	atcgtaaccg	gcgggaccca	gggtatcggt	60
ttggcaatcg	ccgataaatt	tgtagaggag	ggtgcgaaag	tagttatcac	cggtcgccgt	120
gcagatgtag	gtgaaaaggc	cgccaaatca	atcggcggta	ctgatgttat	tcgctttgtc	180
cagcacgatg	catccgatga	agcaggctgg	acgaaactgt	tcgacaccac	cgaggaggca	240
ttcggcccg	ttacgaccgt	cgtgaacaat	gcaggggcct	ctctgaacaa	aagcgttgaa	300
gacactacca	cggaggaatg	gcgcaaactg	ttgtccgtta	atctggatag	tgtttttttc	360
ggcaccgcgc	tgggcattcg	ccgcatgaag	aataaaggct	tgggcgctag	catcatcaat	420
atgagctcga	tcagtgggct	ggtaggcgat	ccgatgttgg	gggcatacaa	cgcttccaag	480
ggggcggtac	gtatcatgtc	gaaaagcgca	gcgctggatt	gcgcagtga	ggactacgat	540
gtgcgtgtca	acacagtaca	tccgggcgct	atcaagaccc	cgctgactga	taaatttcca	600
attggtgggg	aaatgatctc	acagcgtacg	agaacccta	tgggccacat	tgggtgaaccg	660
aatgacgtgg	catggatctg	tgtgtacctg	gcctctgacg	aatcgaaatt	tgcgacgggt	720
agcgaatttg	tggtcgacgg	cgggtatacc	gcacagtga			759

35

<210> 90
 <211> 252
 <212> PRT
 <213> Artificial Sequence

40

<220>
 <223> Variant of L. kefir ketoreductase

<400> 90

45

50

55

EP 2 566 497 B1

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

5 Gln Gly Ile Gly Leu Ala Ile Ala Asp Lys Phe Val Glu Glu Gly Ala
20 25 30

10 Lys Val Val Ile Thr Gly Arg Arg Ala Asp Val Gly Glu Lys Ala Ala
35 40 45

15 Lys Ser Ile Gly Gly Thr Asp Val Ile Arg Phe Val Gln His Asp Ala
50 55 60

Ser Asp Glu Ala Gly Trp Thr Lys Leu Phe Asp Thr Thr Glu Glu Ala
65 70 75 80

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

25 Lys Ser Val Glu Asp Thr Thr Thr Glu Glu Trp Arg Lys Leu Leu Ser
100 105 110

30 Val Asn Leu Asp Ser Val Phe Phe Gly Thr Arg Leu Gly Ile Arg Arg
115 120 125

35

40

45

50

55

EP 2 566 497 B1

Met Lys Asn Lys Gly Leu Gly Ala Ser Ile Ile Asn Met Ser Ser Ile
130 135 140

5 Ser Gly Leu Val Gly Asp Pro Met Leu Gly Ala Tyr Asn Ala Ser Lys
145 150 155 160

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

15 Lys Asp Tyr Asp Val Arg Val Asn Thr Val His Pro Gly Ala Ile Lys
180 185 190

Thr Pro Leu Thr Asp Lys Phe Pro Ile Gly Gly Glu Met Ile Ser Gln
195 200 205

20 Arg Thr Arg Thr Pro Met Gly His Ile Gly Glu Pro Asn Asp Val Ala
210 215 220

25 Trp Ile Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly
225 230 235 240

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

30

<210> 91

<211> 759

<212> DNA

<213> Artificial Sequence

35

<220>

<223> Variant of L. kefir ketoreductase

40

<400> 91

45

50

55

EP 2 566 497 B1

	atgaccaacc gtctgaagag caaagtagcc atcgtaaccg gcgggaccca gggatatcgg	60
	ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttatcac cggtcgccgt	120
5	gcagatgtag gtgaaaaggc cgccaaatca atcggcggta ctgatgttat tcgctttgtc	180
	cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca	240
	ttcggcccgg ttacgaccgt cgtgaacaat gcaggggcct ctctgaacaa aagcgttgaa	300
10	gacactacca cggaggaatg gcgcaaactg ttgtccgtta atctggatag tgtttttttc	360
	ggcacccgtc tgggcattcg ccgcatgaag aataaaggct tgggcgctag catcatcaat	420
	atgagctcga tcagtgggct ggtaggcgat ccgatgttg gggcatacaa cgcttccaag	480
15	ggggcgggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat	540
	gtgcgtgtca acacagtaca tccgggcgct atcaagacc cgctgactga taaatttcca	600
20	ttgggtgggg aaatgatctc acagcgtacg agaacccta tgggccacat tgggtgaaccg	660
	aatgacgtgg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt	720
25	agcgaatttg tggtcgacgg cgggtatacc gcacagtga	759

<210> 92

<211> 252

<212> PRT

30 <213> Artificial Sequence

<220>

<223> Variant of L. kefir ketoreductase

35 <400> 92

40

45

50

55

EP 2 566 497 B1

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

EP 2 566 497 B1

<210> 93
 <211> 759
 <212> DNA
 <213> Artificial Sequence

5

<220>
 <223> Variant of L. kefir ketoreductase

<400> 93

10

```

atgaccaacc gtctgaagag caaagtagcc atcgtaaccg gcgggaccca gggatatcgg      60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttatcac cggtcgcccg      120
gcagatgtag gtgaaaaggc cgccaaatca atcggcggtta ctgatgttat tcgctttgtc      180
cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca      240
ttcggcccggtt tacgaccgt cgtgaacaat gcaggggcct ctctgaacaa aagcgttgaa      300
gacactacca cggaggaatg gcgcaaactg ttgtccgtta atctggatag tgtttttttc      360
ggcaccgcgtc tgggcattcg ccgcatgaag aataaaggct tgggcgctag catcatcaat      420
atgagctcga tcagtgggct ggtaggcgat ccgatgttgg gggcatacaa cgcttccaag      480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat      540
gtgcgtgtca acacagtaca tccgggcgct atcaagacc cgctgactga taaatttcca      600
gctggtgggg aaatgatctc acagcgtacg agaacccta tgggccacat tggatgaacc      660
aatgacgtgg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt      720
agcgaatttg tggtcgacgg cgggtatacc gcacagtga                                759
    
```

35

<210> 94
 <211> 252
 <212> PRT
 <213> Artificial Sequence

40

<220>
 <223> Variant of L. kefir ketoreductase

<400> 94

45

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

50

Gln Gly Ile Gly Leu Ala Ile Ala Asp Lys Phe Val Glu Glu Gly Ala
 20 25 30

55

EP 2 566 497 B1

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

<210> 95

<211> 759

55 <212> DNA

<213> Artificial Sequence

<220>

EP 2 566 497 B1

<223> Variant of L. kefir ketoreductase

<400> 95

5	atgaccaacc gtctgaagag caaagtagcc atcgtaaccg gcgggaccca gggatatcgg	60
	ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttatcac cggtcgccgt	120
	gcagatgtag gtgaaaaggc cgccaaatca atcggcggta ctgatgttat tcgttttgtc	180
10	cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca	240
	ttcggcccggt ttacgaccgt cgtgaacaat gcaggggcct ctgtgaacaa aagcgttgaa	300
15	gacactacca cggaggaatg gcgcaaactg ttgtccgtta atctggatag tgtttttttc	360
	ggcaccocgtc tgggcattca gcgcatgaag aataaaggct tgggcgctag catcatcaat	420
	atgagctcga tcagtgggct ggtaggcgat ccgatgttgg gggcatacaa cgcttccaag	480
20	ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat	540
	gtgcgtgtca acacagtaca tccgggcgct atcaagacc cgctgactga taaatttccg	600
	ggtggtgggg aaatgatctc acagcgtacg aaaacccta tgggccacat tggngaaccg	660
25	aatgacgtgg catggatctg tgtgtacctg gcacttgacg aatcgaaatt tgcgacgggt	720
	agcgaatttg tggtcgacgg cgggtatacc gcacagtga	759

30 <210> 96
 <211> 252
 <212> PRT
 <213> Artificial Sequence

35 <220>
 <223> Variant of L. kefir ketoreductase
 <400> 96

40

45

50

55

EP 2 566 497 B1

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

EP 2 566 497 B1

<210> 97
 <211> 759
 <212> DNA
 <213> Artificial Sequence

5

<220>
 <223> Variant of L. kefir ketoreductase

<400> 97

10

atgaccaacc	gtctgaagag	caaagtagcc	atcgtaaccg	gcgggaccca	gggtatcggt	60
ttggcaatcg	ccgataaatt	tgtagaggag	ggtgcgaaag	tagttatcac	cggtcgccgt	120
gcagatgtag	gtgaaaaggc	cgccaaatca	atcggcggta	ctgatgttat	tcgctttgtc	180
cagcacgatg	catccgatga	agcaggctgg	acgaaactgt	tcgacaccac	cgaggaggca	240
ttcggcccg	ttacgaccgt	cgtgaacaat	gcagggatct	ctatgaacaa	aagcgttgaa	300
gacactacca	cggaggaatg	gcgcaaactg	ttgtccgtta	atctggatag	tgtttttttc	360
ggcaccocgtc	tgggcattcg	ccgcatgaag	aataaaggct	tgggcgctag	catcatcaat	420
atgagctcga	tcagtgggat	ggtaggcgat	ccgatgttgg	gggcatacaa	cgcttccaag	480
ggggcggtac	gtatcatgtc	gaaaagcgca	gcgctggatt	gcgcagtga	ggactacgat	540
gtgcgtgtca	acacagtaca	tccgggcgct	atcaagaccc	cgctgactga	taaatttcca	600
attggtgggg	aaatgatctc	acagcgtacg	agaacccta	tgggccacat	tggatgaaccg	660
aatgacgtgg	catggatctg	tgtgtacctg	gcattctgacg	aatcgaaatt	tgcgacgggt	720
agcgaatttg	tggtcgacgg	cgggtatacc	gcacagtga			759

35

<210> 98
 <211> 252
 <212> PRT
 <213> Artificial Sequence

40

<220>
 <223> Variant of L. kefir ketoreductase

<400> 98

45

50

55

EP 2 566 497 B1

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

EP 2 566 497 B1

<210> 99
 <211> 759
 <212> DNA
 <213> Artificial Sequence

5

<220>
 <223> Variant of L. kefir ketoreductase

<400> 99

10

atgaccaacc	gtctgaagag	caaagtagcc	atcgtaaccg	gcgggaccca	gggtatcggt	60
ttggcaatcg	ccgataaatt	tgtagaggag	ggtgcgaaag	tagttatcac	cggtcgccgt	120
gcagatgtag	gtgaaaaggc	cgccaaatca	atcggcggtg	ctgatgttat	tcgctttgtc	180
cagcacgatg	catccgatga	agcaggctgg	acgaaactgt	tcgacaccac	cgaggaggca	240
ttcggcccgg	ttacgaccgt	cgtgaacaat	gcagggatct	ctctgcctaa	aagcgttgaa	300
gacactacca	cggaggaatg	gcgcaaactg	ttgtccgtta	atctggatag	tgtttttttc	360
ggcaccgcgc	tgggcattcg	ccgcatgaag	aataaaggct	tgggcgctag	catcatcaat	420
atgagctcga	tcagtgggct	ggtaggcgat	ccgatgttgg	gggcatacaa	cgcttccaag	480
ggggcggtac	gtatcatgtc	gaaaagcgca	gcgctggatt	gcgcagtga	ggactacgat	540
gtgcgtgtca	acacagtaca	tccgggcgct	atcaagacct	cgctgactga	taaatttcca	600
gcaggtgggg	aaatgatctc	acagcgtacg	aaaaccctta	tgggccacat	tggtgaaccg	660
aatgacgtgg	catggatctg	tgtgtacctg	gcctctgacg	aatcgaaatt	tgcgacgggt	720
agcgaatttg	tggtcgacgg	cgggtatacc	gcacagtga			759

35

<210> 100
 <211> 252
 <212> PRT
 <213> Artificial Sequence

40

<220>
 <223> Variant of L. kefir ketoreductase

<400> 100

45

Met Thr Asn Arg Leu Lys Ser Lys Val Ala Ile Val Thr Gly Gly Thr

50

55

EP 2 566 497 B1

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

EP 2 566 497 B1

<210> 101
 <211> 759
 <212> DNA
 <213> Artificial Sequence

5

<220>
 <223> Variant of L. kefir ketoreductase

<400> 101

10

atgaccaacc	gtctgaagag	caaagtagcc	atcgtaaccc	gcgggaccca	gggtatcggt	60
tttgcaatcg	cccgcaaat	tgtagaggag	ggtgcgaaag	tagttatcac	cggtcgccgt	120
gcagatgtag	gtgaaaaggc	cgccaaatca	atcggcggtta	ctgatgttat	tcgctttgtc	180
cagcacgatg	catccgatga	agcaggctgg	acgaaactgt	tcgacaccac	cgaggaggca	240
ttcggcccgg	ttacgaccgt	cgtgaacaat	gcaggggcct	ctctgcctaa	aagcgttgaa	300
gacactacca	cggaggaatg	gaaaaaactg	ttgtccgtta	atctggatgg	cgtttttttc	360
ggcaccgcgc	tgggcattcg	ccgcatgaag	aataaaggct	tgggcgctag	catcatcaat	420
atgagctcga	tcagtgggat	ggtaggcgat	ccgatgttgg	gggcatacaa	cgcttccaag	480
ggggcggtac	gtatcatgtc	gaaaagcgca	gcgctggatt	gcgcagtga	ggactacgat	540
gtgcgtgtca	acacagtaca	tccgggcgct	atcaagacc	cgctgactga	taaatttcca	600
ggtggtgggg	aaatgatctc	acagcgtagc	agaacccta	tgggccacat	tggatgaaccg	660
gatgacgtgg	catggatctg	tgtgtacctg	gcacgtgacg	aatcgaaatt	tgcgacgggt	720
agcgaatttg	tggtcgacgg	cgggtatacc	gcacagtga			759

35

<210> 102
 <211> 252
 <212> PRT
 <213> Artificial Sequence

40

<220>
 <223> Variant of L. kefir ketoreductase

<400> 102

45

50

55

EP 2 566 497 B1

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

25

30

35

40

45

50

55

EP 2 566 497 B1

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

<210> 103

<211> 759

<212> DNA

<213> Artificial Sequence

<220>

<223> Variant of L. kefir ketoreductase

<400> 103

EP 2 566 497 B1

	atgaccaacc gtctgaagag caaagtagcc atcgtaaccg gcgggaccca gggatatcgg	60
	cgcgcaatcg cccgcaaatt tgtagaggag ggtgcgaaag tagttatcac cggtcgccgt	120
5	gcagatgtag gtgaaaaggc cgccaaatca atcggcggtta ctgatgttat tcgctttgtc	180
	cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca	240
	ttcggcccggtt ttacgaccgt cgtgaacaat gcaggggcct ctctgcctaa aagcgttgaa	300
10	gacactacca cggaggaatg gaaaaaactg ttgtccgtta atctggatgg cgtttttttc	360
	ggcaccgcgtc tgggcattcg ccgcatgaag aataaaggct tgggcgctag catcatcaat	420
15	atgagctoga tcagtgggat ggtaggcgat ccgatgttgg gggcatacaa cgcttccaag	480
	ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat	540
	gtgcgtgtca acacagtaca tccgggcgct atcaagacc cgctgactga taaatttcca	600
20	ggtggtgggg aaatgatctc acagcgtagc agaacccta tgggccacat tggatgaaccg	660
	gatgacgtgg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt	720
25	agcgaatttg tggtcgacgg cgggtatacc gcacagtga	759

<210> 104

<211> 252

<212> PRT

30 <213> Artificial Sequence

<220>

<223> Variant of L. kefir ketoreductase

35 <400> 104

40

45

50

55

EP 2 566 497 B1

	Met	Thr	Asn	Arg	Leu	Lys	Ser	Lys	Val	Ala	Ile	Val	Thr	Gly	Gly	Thr	
	1				5					10					15		
5	Gln	Gly	Ile	Gly	Arg	Ala	Ile	Ala	Arg	Lys	Phe	Val	Glu	Glu	Gly	Ala	
				20					25					30			
10	Lys	Val	Val	Ile	Thr	Gly	Arg	Arg	Ala	Asp	Val	Gly	Glu	Lys	Ala	Ala	
			35					40					45				
15	Lys	Ser	Ile	Gly	Gly	Thr	Asp	Val	Ile	Arg	Phe	Val	Gln	His	Asp	Ala	
		50					55					60					
20	Ser	Asp	Glu	Ala	Gly	Trp	Thr	Lys	Leu	Phe	Asp	Thr	Thr	Glu	Glu	Ala	
	65					70					75					80	
25	Phe	Gly	Pro	Val	Thr	Thr	Val	Val	Asn	Asn	Ala	Gly	Ala	Ser	Leu	Pro	
					85					90					95		
30	Lys	Ser	Val	Glu	Asp	Thr	Thr	Thr	Glu	Glu	Trp	Lys	Lys	Leu	Leu	Ser	
				100					105					110			
35	Val	Asn	Leu	Asp	Gly	Val	Phe	Phe	Gly	Thr	Arg	Leu	Gly	Ile	Arg	Arg	
			115					120					125				
40	Met	Lys	Asn	Lys	Gly	Leu	Gly	Ala	Ser	Ile	Ile	Asn	Met	Ser	Ser	Ile	
		130					135					140					
45	Ser	Gly	Met	Val	Gly	Asp	Pro	Met	Leu	Gly	Ala	Tyr	Asn	Ala	Ser	Lys	
	145					150					155					160	
50	Gly	Ala	Val	Arg	Ile	Met	Ser	Lys	Ser	Ala	Ala	Leu	Asp	Cys	Ala	Val	
55																	

EP 2 566 497 B1

		165		170		175	
5	Lys Asp Tyr Asp Val Arg Val Asn Thr Val His Pro Gly Ala Ile Lys	180	185	190			
10	Thr Pro Leu Thr Asp Lys Phe Pro Gly Gly Gly Glu Met Ile Ser Gln	195	200	205			
15	Arg Thr Arg Thr Pro Met Gly His Ile Gly Glu Pro Asp Asp Val Ala	210	215	220			
20	Trp Ile Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly	225	230	235			
25	Ser Glu Phe Val Val Asp Gly Gly Tyr Thr Ala Gln	245	250				
30	<210> 105 <211> 759 <212> DNA <213> Artificial Sequence <220> <223> Variant of L. kefir ketoreductase <400> 105						
35	atgaccaacc gtctgaagag caaagtagcc atcgtaaccg gcgggaccca ggttatcggg					60	
40	tttgcaatcg cccgcaaatt tgtagaggag ggtgcgaaag tagttatcac cggtcgccgt					120	
45	gcagatgtag gtgaaaaggc cgccaaatca atcggcggta ctgatgttat tcgctttgtc					180	
50	cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca					240	
55	ttcggccccg ttacgaccgt cgtgaacaat gcaggggcct ctctgcctaa aagcgttgaa					300	
	gacactacca cggaggaatg gaaaaaactg ttgtccgtta atctggatag tggttttttc					360	
	ggcacccgtc tgggcattaa gcgcatgaag aataaaggct tgggcgctag catcatcaat					420	
	atgagctcga tcagtgggat ggtaggcgat ccgatgttgg gggcatacaa cgcttccaag					480	
	ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat					540	
	gtgcgtgtca acacagtaca tccgggcgct atcaagaccc cgctgactga taaatttcca					600	
	ggtggtgggg aaatgatctc acagcgtacg agaacccta tgggccacat tgggtgaaccg					660	
	gatgacgtgg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt					720	
	agcgaatttg tggtcgacgg cgggtatacc gcacagtga					759	
	<210> 106 <211> 252						

EP 2 566 497 B1

<212> PRT

<213> Artificial Sequence

<220>

5 <223> Variant of L. kefir ketoreductase

<400> 106

10

15

20

25

30

35

40

45

50

55

EP 2 566 497 B1

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

245

250

5 <210> 107
 <211> 759
 <212> DNA
 <213> Artificial Sequence

 10 <220>
 <223> Variant of L. kefir ketoreductase

 <400> 107

 15 atgaccaacc gtctgaagag caaagtagcc atcgtaaccg gcgggaccca gggatatcggt 60
 cgcgcaatcg cccgcaaatt tgtagaggag ggtgcgaaag tagttatcac cggtcgccgt 120
 gcagatgtag gtgaaaaggc cgccaaatca atcggcggtta ctgatgttat tcgctttgtc 180
 20 cagcacgatg catccgatga agcaggetgg acgaaactgt tcgacaccac cgaggaggca 240
 ttccggcccg ttacgaccgt cgtgaacaat gcaggggcct ctctgcctaa aagcgttgaa 300
 gacactacca cggaggaatg gaaaaaactg ttgtccgtta atctggatgg cgtttttttc 360
 25 ggcaccogtc tgggcattcg ccgcatgaag aataaaggct tgggcgctag catcatcaat 420
 atgagctoga tcagtgggat ggtaggcgat ccgatgttgg gggcatacaa cgcttccaag 480
 30 ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat 540
 gtgcgtgtca acacagtaca tccgggcgct atcaagacc cgctgactga taaatttcca 600
 ggtggtgggg aaatgatctc acagcgtacg agaacccta tgggccacat tgggaaccg 660
 35 aatgacgtgg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
 agcgaatttg tggtcgacgg cgggtatacc gcacagtga 759

40 <210> 108
 <211> 252
 <212> PRT
 <213> Artificial Sequence

45 <220>
 <223> Variant of L. kefir ketoreductase

 <400> 108

50

55

EP 2 566 497 B1

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

EP 2 566 497 B1

<210> 109
 <211> 759
 <212> DNA
 <213> Artificial Sequence

5

<220>
 <223> Variant of L. kefir ketoreductase

<400> 109

10

atgaccaacc	gtctgaagag	caaagtagcc	atcgtaaccg	gcgggaccca	gggtatcggt	60
tttgcaatcg	cccgcaaatt	tgtagaggag	ggtgcgaaag	tagttatcac	cggtcgccgt	120
gcagatgtag	gtgaaaaggc	cgccaaatca	atcggcggtg	ctgatgttat	tcgctttgtc	180
cagcacgatg	catccgatga	agcaggctgg	acgaaactgt	tcgacaccac	cgaggaggca	240
ttcggcccg	ttacgaccgt	cgtgaacaat	gcaggggcct	ctctgcctaa	aagcgttgaa	300
gacactacca	cggaggaatg	gaaaaaactg	ttgtccgtta	atctggatgg	cgtttttttc	360
ggcacccgtc	tgggcattcg	ccgcatgaag	aataaaggct	tgggcgctag	catcatcaat	420
atgagctcga	tcagtgggat	ggtaggcgat	ccgatgttgg	gggcatacaa	cgtttccaag	480
ggggcgatcc	gtatcatgtc	gaaaagcgca	gcgctggatt	gcgcagtga	ggactacgat	540
gtgcgtgtca	acacagtaca	tccgggcgct	atcaagacct	cgctgactga	taaattttcca	600
ggtggtgggg	aaatgatctc	acagcgtacg	agaaccctta	tgggccacat	tggtgaaccg	660
gatgacgtgg	catggatctg	tgtgtacctg	gcattctgacg	aatcgaaatt	tgcgacgggt	720
agcgaatttg	tggtcgacgg	cgggtatacc	gcacagtga			759

35

<210> 110
 <211> 252
 <212> PRT
 <213> Artificial Sequence

40

<220>
 <223> Variant of L. kefir ketoreductase

<400> 110

45

50

55

EP 2 566 497 B1

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

EP 2 566 497 B1

<210> 111
 <211> 759
 <212> DNA
 <213> Artificial Sequence

5

<220>
 <223> Variant of L. kefir ketoreductase

<400> 111

10

```

atgaccaacc gtctgaagag caaagtagcc atcgtaaccg gcgggaccca gggatatcgg      60
cgcgcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttatcac cggtcgcccg      120
gcagatgtag gtgaaaaggc cgccaaatca atcggcggtta ctgatgttat tcgctttgtc      180
cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca      240
ttcggcccgg ttacgaccgt cgtgaacaat gcaggggcct ctctgcctaa aagcgttgaa      300
gacactacca cggaggaatg gcgcaaactg ttgtccgta atctggatgg cgtttttttc      360
ggcaccgcgc tgggcattcg ccgcatgaag aataaaggct tgggcgctag catcatcaat      420
atgagctcga tcagtgggat ggtaggcgat ccgatgttgg gggcatacaa cgcttccaag      480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat      540
gtgcgtgtca acacagtaca tccgggcgct atcaagacc cgctgactga taaatttcca      600
gggtggtggg aaatgatctc acagcgtacg agaacccta tgggccacat tggngaaccg      660
gatgacgtgg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt      720
agcgaatttg tggtcgacgg cgggtatacc gcacagtga      759
  
```

35

<210> 112
 <211> 252
 <212> PRT
 <213> Artificial Sequence

40

<220>
 <223> Variant of L. kefir ketoreductase

<400> 112

45

50

55

EP 2 566 497 B1

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

EP 2 566 497 B1

<210> 113
<211> 759
<212> DNA
<213> Artificial Sequence

5

<220>
<223> Variant of L. kefir ketoreductase

<400> 113

10

```

atgaccaacc gtctgaagag caaagtagcc atcgtaaccc gcgggaccca gggatatcgg      60
cgcgcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttatcac cggtcgccgt      120
gcagatgtag gtgaaaaggc cgccaaatca atcggcggtta ctgatgttat tcgctttgtc      180
cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca      240
ttcggcccggt ttacgaccgt cgtgaacaat gcaggggcct ctctgcctaa aagcgttgaa      300
gacactacca cggaggaatg gcgcaaactg ttgtccgcta atctggatgg tgtttttttc      360
ggcaccgcgc tgggcattaa gcgcatgaag aataaaggct tgggcgctag catcatcaat      420
atgagctcga tcagtgggat ggtaggcgat ccgatgttgg gggcatacaa cgcttccaag      480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat      540
gtgcgtgtca acacagtaca tccgggcgct atcaagacc cgctgactga taaatttcca      600
gggtggtgggg aaatgatctc acagcgtacg agaacccta tgggccacat tggngaaccg      660
gatgacgtgg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt      720
agcgaatttg tggtcgacgg cgggtatacc gcacagtga      759

```

35

<210> 114
<211> 252
<212> PRT
<213> Artificial Sequence

40

<220>
<223> Variant of L. kefir ketoreductase

<400> 114

45

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

50

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

55

Lys Val Val Ile Thr Gly Arg Arg Ala Asp Val Gly Glu Lys Ala Ala
35 40 45

EP 2 566 497 B1

Lys Ser Ile Gly Gly Thr Asp Val Ile Arg Phe Val Gln His Asp Ala
50 55 60

5 Ser Asp Glu Ala Gly Trp Thr Lys Leu Phe Asp Thr Thr Glu Glu Ala
65 70 75 80

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

Lys Ser Val Glu Asp Thr Thr Thr Glu Glu Trp Arg Lys Leu Leu Ser
100 105 110

15 Val Asn Leu Asp Gly Val Phe Phe Gly Thr Arg Leu Gly Ile Lys Arg
115 120 125

20 Met Lys Asn Lys Gly Leu Gly Ala Ser Ile Ile Asn Met Ser Ser Ile
130 135 140

25 Ser Gly Met Val Gly Asp Pro Met Leu Gly Ala Tyr Asn Ala Ser Lys
145 150 155 160

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

30 Lys Asp Tyr Asp Val Arg Val Asn Thr Val His Pro Gly Ala Ile Lys
180 185 190

35 Thr Pro Leu Thr Asp Lys Phe Pro Gly Gly Gly Glu Met Ile Ser Gln
195 200 205

40 Arg Thr Arg Thr Pro Met Gly His Ile Gly Glu Pro Asp Asp Val Ala
210 215 220

Trp Ile Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly
225 230 235 240

45 Ser Glu Phe Val Val Asp Gly Gly Tyr Thr Ala Gln
245 250

<210> 115

50 <211> 759

<212> DNA

<213> Artificial Sequence

<220>

55 <223> Variant of L. kefir ketoreductase

<400> 115

atgaccaacc gtctgaagag caaagtagcc atcgtaaccg gcgggaccca gggatatcgg 60
 ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttatcac cggtcgccgt 120
 5 gcagatgtag gtgaaaaggc cgccaaatca atcggcggtta ctgatgttat tcgctttgtc 180
 cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
 10 ttcgggcccg ttacgaccgt cgtgaacaat gcaggggcct ctctgcctaa aagcgttgaa 300
 gacactacca cggaggaatg gcgcaaactg ttgtccgtta atctggatgg cgtttttttc 360
 ggcacccgtc tgggcattcg ccgcatgaag aataaaggct tgggcgctag catcatcaat 420
 15 atgagctcga tcagtgggat ggtaggcgat ccgatgttgg gggcatacaa cgcttccaag 480
 ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat 540
 20 gtgcgtgtca acacagtaca tccgggcgct atcaagaccc cgctgactga taaatttcca 600
 ggtggtgggg aaatgatctc acagcgtacg agaacccta tgggccacat tggatgaaccg 660
 gatgacgtgg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
 25 agcgaatttg tggtcgacgg cgggtatacc gcacagtga 759

<210> 116

<211> 252

<212> PRT

30 <213> Artificial Sequence

<220>

<223> Variant of L. kefir ketoreductase

35 <400> 116

40

45

50

55

EP 2 566 497 B1

	Met	Thr	Asn	Arg	Leu	Lys	Ser	Lys	Val	Ala	Ile	Val	Thr	Gly	Gly	Thr
	1				5					10					15	
5	Gln	Gly	Ile	Gly	Leu	Ala	Ile	Ala	Asp	Lys	Phe	Val	Glu	Glu	Gly	Ala
				20					25					30		
10	Lys	Val	Val	Ile	Thr	Gly	Arg	Arg	Ala	Asp	Val	Gly	Glu	Lys	Ala	Ala
			35					40					45			
15	Lys	Ser	Ile	Gly	Gly	Thr	Asp	Val	Ile	Arg	Phe	Val	Gln	His	Asp	Ala
		50					55					60				
20	Ser	Asp	Glu	Ala	Gly	Trp	Thr	Lys	Leu	Phe	Asp	Thr	Thr	Glu	Glu	Ala
	65					70					75					80
25	Phe	Gly	Pro	Val	Thr	Thr	Val	Val	Asn	Asn	Ala	Gly	Ala	Ser	Leu	Pro
					85					90					95	
30	Lys	Ser	Val	Glu	Asp	Thr	Thr	Thr	Glu	Glu	Trp	Arg	Lys	Leu	Leu	Ser
				100					105					110		
35	Val	Asn	Leu	Asp	Gly	Val	Phe	Phe	Gly	Thr	Arg	Leu	Gly	Ile	Arg	Arg
			115					120					125			

EP 2 566 497 B1

	Met	Lys	Asn	Lys	Gly	Leu	Gly	Ala	Ser	Ile	Ile	Asn	Met	Ser	Ser	Ile
	130						135					140				
5	Ser	Gly	Met	Val	Gly	Asp	Pro	Met	Leu	Gly	Ala	Tyr	Asn	Ala	Ser	Lys
	145					150					155					160
10	Gly	Ala	Val	Arg	Ile	Met	Ser	Lys	Ser	Ala	Ala	Leu	Asp	Cys	Ala	Val
					165					170					175	
15	Lys	Asp	Tyr	Asp	Val	Arg	Val	Asn	Thr	Val	His	Pro	Gly	Ala	Ile	Lys
				180					185					190		
20	Thr	Pro	Leu	Thr	Asp	Lys	Phe	Pro	Gly	Gly	Gly	Glu	Met	Ile	Ser	Gln
			195					200					205			
25	Arg	Thr	Arg	Thr	Pro	Met	Gly	His	Ile	Gly	Glu	Pro	Asp	Asp	Val	Ala
	210						215					220				
30	Trp	Ile	Cys	Val	Tyr	Leu	Ala	Ser	Asp	Glu	Ser	Lys	Phe	Ala	Thr	Gly
	225					230					235					240
35	Ser	Glu	Phe	Val	Val	Asp	Gly	Gly	Tyr	Thr	Ala	Gln				
					245					250						

<210> 117

<211> 759

<212> DNA

<213> Artificial Sequence

<220>

<223> Variant of L. kefir ketoreductase

<400> 117

atgaccaacc gtctgaagag caaagtagcc atcgtaaccg gcgggaacca gggatatcgg 60
 ttagcaatcg cccgcaaatt tgtagaggag ggtgcgaaag tagttatcac cggtcgccgt 120
 5 gcagatgtag gtgaaaaggc cgccaaatca atcggcggta ctgatgttat tcgctttgtc 180
 cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
 ttccggcccg ttacgaccgt cgtgaacaat gcaggggcct ctctgcctaa aagcgttgaa 300
 10 gacactacca cggaggaatg ggacaaactg ttgtccgtta atctggatgg cgtttttttc 360
 ggcacccgtc tgggcattaa gcgcatgaag aataaaggct tgggcgctag catcatcaat 420
 atgagctcga tcagtgggat ggtaggcgat ccgatgttgg gggcatacaa cgcttccaag 480
 15 ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat 540
 gtgcgtgtca acacagtaca tccgggcgct atcaagacc cgctgactga taaatttcca 600
 20 ggtggtgggg aaatgatctc acagcgtacg agaacccta tgggccacat tgggtgaaccg 660
 gatgacgtgg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
 25 agcgaatttg tggtcgacgg cgggtatacc gcacagtga 759

<210> 118

<211> 252

<212> PRT

30 <213> Artificial Sequence

<220>

<223> Variant of L. kefir ketoreductase

35 <400> 118

40

45

50

55

EP 2 566 497 B1

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

EP 2 566 497 B1

<210> 119
 <211> 759
 <212> DNA
 <213> Artificial Sequence

5

<220>
 <223> Variant of L. kefir ketoreductase

10

```

atgaccaacc gtctgaagag caaagtagcc atcgtaaccg gcgggaccca gggatatcgg      60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttatcac cggtcgccgt      120
gcagatgtag gtgaaaaggc cgccaaatca atcggcggtta ctgatgttat tcgctttgtc      180
cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca      240
ttcggcccgg ttacgaccgt cgtgaacaat gcaggggcct ctctgcctaa aagcgttgaa      300
gacactacca cggaggaatg gcgcaaactg ttgtccgta atctggatgg cgtttttttc      360
ggcaccgcgc tgggcattcg ccgcatgaag aataaaggct tgggcgctag catcatcaat      420
atgagctcga tcagtgggat ggtaggcgat ccgatgttgg gggcatacaa cgcttccaag      480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat      540
gtgcgtgtca acacagtaca tccgggcgct atcaagacc cgctgactga taaatttcca      600
ggtggtgggg aaatgatctc acagcgtacg agaacccta tgggccacat tggngaaccg      660
aatgacgtgg catggatctg tgtgtacctg gcactctgac aatcgaaatt tgcgacgggt      720
agcgaatttg tggtcgacgg cgggtatacc gcacagtga      759
  
```

35

<210> 120
 <211> 252
 <212> PRT
 <213> Artificial Sequence

40

<220>
 <223> Variant of L. kefir ketoreductase

45

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

50

Gln Gly Ile Gly Leu Ala Ile Ala Asp Lys Phe Val Glu Glu Gly Ala
 20 25 30

55

EP 2 566 497 B1

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

<210> 121

<211> 759

55 <212> DNA

<213> Artificial Sequence

<220>

EP 2 566 497 B1

<223> Variant of L. kefir ketoreductase

<400> 121

```

5  atgaccaacc gtctgaagag caaagtagcc atcgtaaccg gcgggaccca gggatatcgg 60
   ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttatcac cggtcgccgt 120
   gcagatgtag gtgaaaaggc cgccaaatca atcggcggta ctgatgttat tcgctttgtc 180
10  cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
   ttcgggcccg ttacgaccgt cgtgaacaat gcaggggcct ctctgcctaa aagcgttgaa 300
   gacactacca cggaggaatg gaaaaaactg ttgtccgtta atctggatgg cgtttttttc 360
   ggcacccgtc tgggcattaa gcgcatgaag aataaaggct tgggcgctag catcatcaat 420
   atgagctcga tcagtgggat ggtaggcgat ccgatgttgg gggcatacaa cgcttccaag 480
20  ggggcgggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat 540
   gtgcgtgtca acacagtaca tccgggcgct atcaagacc cgctgactga taaatttcca 600
   ggtggtgggg aaatgatctc acagcgtacg agaacccta tgggccacat tggngaaccg 660
25  gatgacgtgg catggatctg tgtgtacctg gcattctgac aatcgaaatt tgcgacgggt 720
   agcgaatttg tggtcgacgg cgggtatacc gcacagtga 759

```

```

30  <210> 122
     <211> 252
     <212> PRT
     <213> Artificial Sequence

```

```

35  <220>
     <223> Variant of L. kefir ketoreductase

     <400> 122

```

40

45

50

55

EP 2 566 497 B1

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

EP 2 566 497 B1

<210> 123
 <211> 759
 <212> DNA
 <213> Artificial Sequence

5

<220>
 <223> Variant of L. kefir ketoreductase

<400> 123

10

atgaccaacc	gtctgaagag	caaagtagcc	atcgtaaccg	gcgggaccca	gggtatcggg	60
cgcgcaatcg	cccgcaaatt	tgtagaggag	ggcgcgaaag	tagttatcac	cggtcgccgt	120
gcagatgtag	gtgaaaaggc	cgccaaatca	atcggcggta	ctgatgttat	tcgctttgtc	180
cagcacgatg	catccgatga	agcaggctgg	acgaaactgt	tcgacaccac	cgaggaggca	240
ttcggccccg	ttacgaccgt	cgtgaacaat	gcaggggcct	ctctgcctaa	aagcgttgaa	300
gacactacca	cggaggaatg	gcgcaaactg	ttgtccgtta	atctggatgg	cgtttttttc	360
ggcaccgcgc	tgggcattaa	gcgcatgaag	aataaaggct	tgggcgctag	catcatcaat	420
atgagctcga	tcagtgggct	ggtaggcgat	ccgatgttgg	gggcatacaa	cgcttccaag	480
ggggcggtac	gtatcatgtc	gaaaagcgca	gcgctggatt	gcgcagtga	ggactacgat	540
gtgcgtgtca	acacagtaca	tccgggcgct	atcaagaccc	cgctgactga	taaattttcca	600
gggtggtggg	aaatgatctc	acagcgtagc	agaacccta	tgggccacat	tgggtgaaccg	660
gatgacgtgg	catggatctg	tgtgtacctg	gcctctgacg	aatcgaaatt	tgcgacgggt	720
agcgaatttg	tggtcgacgg	cggttatacc	gcacagtga			759

30

35

<210> 124
 <211> 252
 <212> PRT
 <213> Artificial Sequence

40

<220>
 <223> Variant of L. kefir ketoreductase

45

<400> 124

50

55

EP 2 566 497 B1

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

EP 2 566 497 B1

<210> 125
 <211> 759
 <212> DNA
 <213> Artificial Sequence

5

<220>
 <223> Variant of L. kefir ketoreductase

<400> 125

10

atgaccaacc	gtctgaagag	caaagtagcc	atcgtaaccg	gcgggaccca	gggtatcggt	60
ttagcaatcg	cccgcaaatt	tgtagaggag	ggtgcgaaag	tagttatcac	cggtcgccgt	120
gcagatgtag	gtgaaaaggc	cgccaaatca	atcggcggtg	ctgatgttat	tcgctttgtc	180
cagcacgatg	catccgatga	agcaggctgg	acgaaactgt	tcgacaccac	cgaggaggca	240
ttcggcccgg	ttacgaccgt	cgtgaacaat	gcaggggcct	ctctgcctaa	aagcgttgaa	300
gacactacca	cggaggaatg	gcgcaaactg	ttgtccgtta	atctggatgg	cgtttttttc	360
ggcaccgcgc	tgggcattaa	gcgcatgaag	aataaaggct	tgggcgctag	catcatcaat	420
atgagctcga	tcagtgggct	ggtaggcgat	ccgatgttgg	gggcatacaa	cgcttccaag	480
ggggcggtac	gtatcatgtc	gaaaagcgca	gcgctggatt	gcgcagtga	ggactacgat	540
gtgcgtgtca	acacagtaca	tccgggcgct	atcaagacc	cgctgactga	taaatttcca	600
ggtggtgggg	aaatgatctc	acagcgtacg	agaaccctta	tgggccacat	tggtgaaccg	660
aatgacgtgg	catggatctg	tgtgtacctg	gcctctgacg	aatcgaaatt	tgcgacgggt	720
agcgaatttg	tggtcgacgg	cgggtatacc	gcacagtga			759

35

<210> 126
 <211> 252
 <212> PRT
 <213> Artificial Sequence

40

<220>
 <223> Variant of L. kefir ketoreductase

<400> 126

45

Met Thr Asn Arg Leu Lys Ser Lys Val Ala Ile Val Thr Gly Gly Thr

50

55

EP 2 566 497 B1

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

EP 2 566 497 B1

<210> 127
 <211> 759
 <212> DNA
 <213> Artificial Sequence

5

<220>
 <223> Variant of L. kefir ketoreductase

<400> 127

10

```

atgaccaacc gtctgaagag caaagtagcc atcgtaaccg gcgggaccca gggatatcgg      60
ttagcaatcg cccgcaaatt tgtagaggag ggtgcgaaag tagttatcac cggtcgccgt      120
gcagatgtag gtgaaaaggc cgccaaatca atcggcggtta ctgatgttat tcgctttgtc      180
cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca      240
ttcggccccg ttacgaccgt cgtgaacaat gcaggggcct ctctgcctaa aagcgttgaa      300
gacactacca cggaggaatg gaaaaaactg ttgtccgta atctggatgg cgtttttttc      360
ggcaccgcgc tgggcattaa gcgcatgaag aataaaggct tgggcgctag catcatcaat      420
atgagctcga tcagtgggat ggtaggcgat ccgatgttgg gggcatacaa cgcttccaag      480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat      540
gtgcgtgtca acacagtaca tccgggcgct atcaagacc cgctgactga taaatttcca      600
gggtggtggg aaatgatctc acagcgtacg agaacccta tgggccacat tgggtgaaccg      660
gatgacgtgg catggatctg tgtgtacctg gcctctgacg aatcgaaatt tgcgacgggt      720
agcgaatttg tggtcgacgg cgggtatacc gcacagtga      759
  
```

35

<210> 128
 <211> 252
 <212> PRT
 <213> Artificial Sequence

40

<220>
 <223> Variant of L. kefir ketoreductase

<400> 128

45

50

55

EP 2 566 497 B1

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

EP 2 566 497 B1

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

<210> 129

<211> 759

<212> DNA

<213> Artificial Sequence

<220>

<223> Variant of L. kefir ketoreductase

<400> 129

EP 2 566 497 B1

	atgaccaacc gtctgaagag caaagtagcc atcgtaaccg gcgggaccca gggatatcgg	60
	ttagcaatcg cccgcaaatt thtagaggag ggtgcgaaag tagttatcac cggtcgccgt	120
5	gcagatgtag gtgaaaaggc cgccaaatca atcggcggtta ctgatgttat tcgctttgtc	180
	cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca	240
	ttcggcccgg ttacgaccgt cgtgaacaat gcaggggcct ctctgcctaa aagcgttgaa	300
10	gacactacca cggaggaatg gcgcaaactg ttgtccgtta atctggatag tgttttttc	360
	ggcaccgcgc tgggcattcg ccgcatgaag aataaaggct tgggcgctag catcatcaat	420
15	atgagctcga tcagtgggat ggtaggcgat ccgatgttg gggcatacaa cgcttccaag	480
	ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat	540
	gtgcgtgtca acacagtaca tccgggcgct atcaagacc cgctgactga taaatttcca	600
20	ggtggtgggg aaatgatctc acagcgtagc agaacccta tgggccacat tggatgaaccg	660
	gatgacgtgg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt	720
25	agcgaatttg tggtcgacgg cgggtatacc gcacagtga	759

<210> 130

<211> 252

<212> PRT

30 <213> Artificial Sequence

<220>

<223> Variant of L. kefir ketoreductase

35 <400> 130

40

45

50

55

EP 2 566 497 B1

	Met	Thr	Asn	Arg	Leu	Lys	Ser	Lys	Val	Ala	Ile	Val	Thr	Gly	Gly	Thr	
	1				5					10					15		
5	Gln	Gly	Ile	Gly	Leu	Ala	Ile	Ala	Arg	Lys	Phe	Val	Glu	Glu	Gly	Ala	
				20					25					30			
10	Lys	Val	Val	Ile	Thr	Gly	Arg	Arg	Ala	Asp	Val	Gly	Glu	Lys	Ala	Ala	
			35					40					45				
	Lys	Ser	Ile	Gly	Gly	Thr	Asp	Val	Ile	Arg	Phe	Val	Gln	His	Asp	Ala	
		50					55					60					
15	Ser	Asp	Glu	Ala	Gly	Trp	Thr	Lys	Leu	Phe	Asp	Thr	Thr	Glu	Glu	Ala	
	65					70					75					80	
20	Phe	Gly	Pro	Val	Thr	Thr	Val	Val	Asn	Asn	Ala	Gly	Ala	Ser	Leu	Pro	
					85					90					95		
25	Lys	Ser	Val	Glu	Asp	Thr	Thr	Thr	Glu	Glu	Trp	Arg	Lys	Leu	Leu	Ser	
				100					105					110			
	Val	Asn	Leu	Asp	Ser	Val	Phe	Phe	Gly	Thr	Arg	Leu	Gly	Ile	Arg	Arg	
			115					120					125				
30	Met	Lys	Asn	Lys	Gly	Leu	Gly	Ala	Ser	Ile	Ile	Asn	Met	Ser	Ser	Ile	
		130					135					140					
35	Ser	Gly	Met	Val	Gly	Asp	Pro	Met	Leu	Gly	Ala	Tyr	Asn	Ala	Ser	Lys	
	145					150					155					160	
40	Gly	Ala	Val	Arg	Ile	Met	Ser	Lys	Ser	Ala	Ala	Leu	Asp	Cys	Ala	Val	

EP 2 566 497 B1

	165	170	175
5	Lys Asp Tyr Asp Val Arg Val Asn Thr Val His Pro Gly Ala Ile Lys 180 185 190		
10	Thr Pro Leu Thr Asp Lys Phe Pro Gly Gly Gly Glu Met Ile Ser Gln 195 200 205		
15	Arg Thr Arg Thr Pro Met Gly His Ile Gly Glu Pro Asp Asp Val Ala 210 215 220		
20	Trp Ile Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly 225 230 235 240		
25	Ser Glu Phe Val Val Asp Gly Gly Tyr Thr Ala Gln 245 250		
30	<210> 131 <211> 759 <212> DNA <213> Artificial Sequence <220> <223> Variant of L. kefir ketoreductase <400> 131		
35	atgaccaacc gtctgaagag caaagtagcc atcgtaaccg gcgggaccca gggatatcgg	60	
40	cgcgcaatcg cccgcaaatt tgtagaggag ggtgcgaaag tagttatcac cggtcgccgt	120	
45	gcagatgtag gtgaaaaggc cgccaaatca atcggcggta ctgatgttat tcgctttgtc	180	
50	cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca	240	
55	ttcggccccg ttacgaccgt cgtgaacaat gcaggggcct ctctgcctaa aagcgttgaa	300	
	gacactacca cggaggaatg gcgcaaactg ttgtccgtta atctggatgg cgtttttttc	360	
	ggcaccgcgtc tgggcattcg ccgcatgaag aataaaggct tgggcgctag catcatcaat	420	
	atgagctcga tcagtgggat ggtaggcgat ccgatgttgg gggcatacaa cgcttccaag	480	
	ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat	540	
	gtgcgtgtca acacagtaca tccgggcgct atcaagaccc cgctgactga taaatttcca	600	
	gggtggtgggg aaatgatctc acagcgtacg agaacccta tgggccacat tgggtgaaccg	660	
	gatgacgtgg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt	720	
	agcgaatttg tggtcgacgg cgggtatacc gcacagtga	759	
	<210> 132 <211> 252		

<212> PRT

<213> Artificial Sequence

<220>

5 <223> Variant of L. kefir ketoreductase

<400> 132

10

15

20

25

30

35

40

45

50

55

EP 2 566 497 B1

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

245

250

5 <210> 133
 <211> 759
 <212> DNA
 <213> Artificial Sequence

10 <220>
 <223> Variant of L. kefir ketoreductase

<400> 133

15 atgaccaacc gtctgaagag caaagtagcc atcgtaaccg gcgggaccca gggatatcgg 60
 ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttatcac cggtcgcccg 120
 gcagatgtag gtgaaaaggc cgccaaatca atcggcggtta ctgatgttat tcgctttgtc 180
 20 cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
 ttccggcccg ttacgaccgt cgtgaacaat gcagggacgt ctctgcctaa aagcgttgaa 300
 gacactacca cggaggaatg gcgcaaactg ttgtccgta atctggatag tgtttttttc 360
 25 ggcaccogtc tgggcattcg ccgcatgaag aataaaggct tgggcgctag catcatcaat 420
 atgagctoga tcagtgggct ggtaggcgat ccgatgttgg gggcatacaa cgcttccaag 480
 30 ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat 540
 gtgcgtgtca acacagtaca tccgggcgct atcaagacct cgctgactga taaatttcca 600
 ggtggtgggg aaatgatctc acagcgtagc agaaccctta tgggccacat tggatgaaccg 660
 35 aatgacgtgg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
 agcgaatttg tggtcgacgg cgggtatacc gcacagtga 759

40 <210> 134
 <211> 252
 <212> PRT
 <213> Artificial Sequence

45 <220>
 <223> Variant of L. kefir ketoreductase

<400> 134

50

55

EP 2 566 497 B1

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

EP 2 566 497 B1

<210> 135
 <211> 759
 <212> DNA
 <213> Artificial Sequence

5

<220>
 <223> Variant of L. kefir ketoreductase

<400> 135

10

atgaccaacc	gtctgaagag	caaagtagcc	atcgtaaccg	gcgggaccca	gggtatcggt	60
ttggcaatcg	ccgataaatt	tgtagaggag	ggtgcgaaag	tagttatcac	cggtcgccgt	120
gcagatgtag	gtgaaaaggc	cgccaaatca	atcggcggtg	ctgatgttat	tcgctttgtc	180
cagcacgatg	catccgatga	agcaggctgg	acgaaactgt	tcgacaccac	cgaggaggca	240
ttcggcccgg	ttacgaccgt	cgtgaacaat	gcagggggccg	cgctgcctaa	aagcgttgaa	300
gacactacca	cggaggaatg	gcgcaaactg	ttgtccgtta	atctggatag	tgtttttttc	360
ggcacccgtc	tgggcattcg	ccgcatgaag	aataaaggct	tgggcgctag	catcatcaat	420
atgagctcga	tcagtgggct	ggtaggcgat	ccgatgttgg	gggcatacaa	cgcttccaag	480
ggggcggtac	gtatcatgtc	gaaaagcgca	gcgctggatt	gcgcagtga	ggactacgat	540
gtgcgtgtca	acacagtaca	tccgggcgct	atcaagaccc	cgctgactga	taaattttcca	600
ggtggtgggg	aaatgatctc	acagcgtagc	agaaccacct	tggggccacat	tgggtgaaccg	660
aatgacgtgg	catggatctg	tgtgtacctg	gcacgtgacg	aatcgaaatt	tgcgacgggt	720
agcgaatttg	tgggtcgacg	cgggtatacc	gcacagtga			759

35

<210> 136
 <211> 252
 <212> PRT
 <213> Artificial Sequence

40

<220>
 <223> Variant of L. kefir ketoreductase

<400> 136

45

50

55

EP 2 566 497 B1

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

EP 2 566 497 B1

<210> 137
 <211> 759
 <212> DNA
 <213> Artificial Sequence

5

<220>
 <223> Variant of L. kefir ketoreductase

<400> 137

10

atgaccaacc	gtctgaagag	caaagtagcc	atcgtaaccg	gcgggaccca	gggtatcggt	60
ttggcaatcg	ccgataaatt	tgtagaggag	ggtgcgaaag	tagttatcac	cggtcgcctt	120
gcagatgtag	gtgaaaaggc	cgccaaatca	atcggcggtg	ctgatgttat	tcgctttgtc	180
cagcacgatg	catccgatga	agcaggctgg	acgaaactgt	tcgacaccac	cgaggaggca	240
ttcggcccg	ttacgaccgt	cgtgaacaat	gcaggggcct	ctctgcctaa	aagcgttgaa	300
gacactacca	cggaggaatg	gcgcaaactg	ttgtccgtta	atctggatag	tgtttttttc	360
ggcaccgcgc	tgggcattcg	ccgcatgaag	aataaaggct	tgggcgctag	catcatcaat	420
atgagctoga	tcagtgggct	ggtaggcgat	ccgatgttgg	gggcatacaa	cgcttccaag	480
ggggcggtac	gtatcatgtc	gaaaagcgca	gcgctggatt	gcgcagtga	ggactacgat	540
gtgcgtgtca	acacagtaca	tccgggcgct	atcaagacc	cgctgactga	taaatttcca	600
ggtggtgggg	aagtgatctc	acagcgtacg	agaaccctta	tgggccacat	tggtgaaccg	660
aatgacgtgg	catggatctg	tgtgtacctg	gcctctgacg	aatcgaaatt	tgcgacgggt	720
agcgaatttg	tggtcgacgg	cggttatacc	gcacagtga			759

35

<210> 138
 <211> 252
 <212> PRT
 <213> Artificial Sequence

40

<220>
 <223> Variant of L. kefir ketoreductase

<400> 138

45

50

55

EP 2 566 497 B1

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

EP 2 566 497 B1

<210> 139
 <211> 759
 <212> DNA
 <213> Artificial Sequence

5

<220>
 <223> Variant of L. kefir ketoreductase

10

```

atgaccaacc gtctgaagag caaagtagcc atcgtaaccg gcgggaccca gggatcggt      60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttatcac cggtcgccgt      120
gcagatgtag gtgaaaaggc cgccaaatca atcggcggtta ctgatgttat tcgctttgtc      180
cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca      240
ttcggcccggt ttacgaccgt cgtgaacaat gcagggatct ctctgcctaa aagcgttgaa      300
gacactacca cggaggaatg gcgcaaactg ttgtccgtta atctggatgc ggtttttttc      360
ggcaccgcgtc tgggcattcg ccgcatgaag aataaaggct tgggcgctag catcatcaat      420
atgagctcga tcagtgggct ggtaggcgat ccgatgttgg gggcatacaa cgcttccaag      480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat      540
gtgcgtgtca acacagtaca tccgggcgct atcaagacc cgctgactga taaatttcca      600
gcaggtgggg aaatgatctc acagcgtacg aaaacccta tgggccacat tggngaaccg      660
aatgacgtgg catggatctg tgtgtacctg gcactctgacg aatcgaaatt tgcgacgggt      720
agcgaatttg tggtcgacgg cgggtatacc gcacagtga      759
  
```

35

<210> 140
 <211> 252
 <212> PRT
 <213> Artificial Sequence

40

<220>
 <223> Variant of L. kefir ketoreductase

45

```

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

Gln Gly Ile Gly Leu Ala Ile Ala Asp Lys Phe Val Glu Glu Gly Ala
          20              25              30

Lys Val Val Ile Thr Gly Arg Arg Ala Asp Val Gly Glu Lys Ala Ala
          35              40              45
  
```

55

EP 2 566 497 B1

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

<210> 141

50 <211> 759

<212> DNA

<213> Artificial Sequence

<220>

55 <223> Variant of L. kefir ketoreductase

<400> 141

atgaccaacc gtctgaagag caaagtagcc atcgtaaccg gcgggaccca gggatatcgg 60
 ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttatcac cggtcgccgt 120
 5 gcagatgtag gtgaaaaggc cgccaaatca atcggcggtta ctgatgttat tcgctttgtc 180
 cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
 10 ttcgggcccg ttacgaccgt cgtgaacaat gcagggatct ctctgcctaa aagcgttgaa 300
 gacactacca cggaggaatg gcgcaaactg ttgtccgtta atctggatgg gggttttttc 360
 ggcaccgcgc tgggcattcg ccgcatgaag aataaaggct tgggcgctag catcatcaat 420
 15 atgagctoga tcagtgggct ggtaggcgat ccgatgttgg gggcatacaa cgcttccaag 480
 ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat 540
 20 gtgcgtgtca acacagtaca tccgggcgct atcaagaccc cgctgactga taaatttcca 600
 gcaggtgggg aaatgatctc acagcgtacg aaaacccta tgggccacat tggatgaaccg 660
 aatgacgtgg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
 25 agcgaatttg tggtcgacgg cgggtatacc gcacagtga 759

<210> 142

<211> 252

<212> PRT

30 <213> Artificial Sequence

<220>

<223> Variant of L. kefir ketoreductase

35 <400> 142

40

45

50

55

EP 2 566 497 B1

	Met	Thr	Asn	Arg	Leu	Lys	Ser	Lys	Val	Ala	Ile	Val	Thr	Gly	Gly	Thr
	1				5					10					15	
5	Gln	Gly	Ile	Gly	Leu	Ala	Ile	Ala	Asp	Lys	Phe	Val	Glu	Glu	Gly	Ala
				20					25					30		
10	Lys	Val	Val	Ile	Thr	Gly	Arg	Arg	Ala	Asp	Val	Gly	Glu	Lys	Ala	Ala
			35					40					45			
15	Lys	Ser	Ile	Gly	Gly	Thr	Asp	Val	Ile	Arg	Phe	Val	Gln	His	Asp	Ala
		50					55					60				
20	Ser	Asp	Glu	Ala	Gly	Trp	Thr	Lys	Leu	Phe	Asp	Thr	Thr	Glu	Glu	Ala
	65					70					75					80
25	Phe	Gly	Pro	Val	Thr	Thr	Val	Val	Asn	Asn	Ala	Gly	Ile	Ser	Leu	Pro
					85					90					95	
30	Lys	Ser	Val	Glu	Asp	Thr	Thr	Thr	Glu	Glu	Trp	Arg	Lys	Leu	Leu	Ser
				100					105					110		
35	Val	Asn	Leu	Asp	Gly	Val	Phe	Phe	Gly	Thr	Arg	Leu	Gly	Ile	Arg	Arg
			115					120					125			
40																
45																
50																
55																

EP 2 566 497 B1

Met Lys Asn Lys Gly Leu Gly Ala Ser Ile Ile Asn Met Ser Ser Ile
130 135 140

5 Ser Gly Leu Val Gly Asp Pro Met Leu Gly Ala Tyr Asn Ala Ser Lys
145 150 155 160

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

15 Lys Asp Tyr Asp Val Arg Val Asn Thr Val His Pro Gly Ala Ile Lys
180 185 190

Thr Pro Leu Thr Asp Lys Phe Pro Ala Gly Gly Glu Met Ile Ser Gln
195 200 205

20 Arg Thr Lys Thr Pro Met Gly His Ile Gly Glu Pro Asn Asp Val Ala
210 215 220

25 Trp Ile Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly
225 230 235 240

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

30

<210> 143

<211> 759

<212> DNA

<213> Artificial Sequence

35

<220>

<223> Variant of L. kefir ketoreductase

40

<400> 143

45

50

55

	atgaccaacc gtctgaagag caaagtagcc atcgtaaccg gcgggaccca gggatatcgg	60
	ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttatcac cggtcgccgt	120
5	gcagatgtag gtgaaaaggc cgccaaatca atcggcggta ctgatgttat tcgctttgtc	180
	cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca	240
	ttcggcccgg ttacgaccgt cgtgaacaat gcagggatct ctctgcctaa aagcgttgaa	300
10	gacactacca cggaggaatg gcgcaaactg ttgtccgtta atctggatag tgtttttttc	360
	ggcaccogtc tgggcattcg ccgcatgaag aataaaggct tgggcgctag catcatcaat	420
	atgagctcga tcagtgggct ggtaggcgat ccgatgttg gggcatacaa cgcttccaag	480
15	ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat	540
	gtgcgtgtca acacagtaca tccgggcgct atcaagacc cgctgactga taaatggcca	600
20	gcaggtgggg aaatgatctc acagcgtacg aaaacccta tgggccacat tggatgaaccg	660
	aatgacgtgg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt	720
25	agcgaatttg tggtcgacgg cgggtatacc gcacagtga	759

<210> 144

<211> 252

<212> PRT

30 <213> Artificial Sequence

<220>

<223> Variant of L. kefir ketoreductase

35 <400> 144

40

45

50

55

EP 2 566 497 B1

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

EP 2 566 497 B1

<210> 145
 <211> 759
 <212> DNA
 <213> Artificial Sequence

5

<220>
 <223> Variant of L. kefir ketoreductase

<400> 145

10

```

atgaccaacc gtctgaagag caaagtagcc atcgtaaccg gcgggaccca gggatatcgg      60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttatcac cggtcgcgct      120
gcagatgtag gtgaaaaggc cgccaaatca atcggcggtta ctgatgttat tcgctttgtc      180
cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca      240
ttcggccccg ttacgaccgt cgtgaacaat gcagggatct ctctgcctaa aagcgttgaa      300
gacactacca cggaggaatg gcgcaaactg ttgtccgta atctggatag tgtttttttc      360
ggcaccgctc tgggcattcg ccgcatgaag aataaaggct tgggcgctag catcatcaat      420
atgagctcga tcagtgggat ggtaggcgat ccgatgttgg gggcatacaa cgcttccaag      480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat      540
gtgcgtgtca acacagtaca tccgggcgct atcaagacc cgctgactga taaatttcca      600
gcaggtgggg aaatgatctc acagcgtacg aaaacccta tgggccacat tggatgaacc      660
aatgacgtgg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt      720
agcgaatttg tggtcgacgg cgggtatacc gcacagtga      759
  
```

35

<210> 146
 <211> 252
 <212> PRT
 <213> Artificial Sequence

40

<220>
 <223> Variant of L. kefir ketoreductase

<400> 146

45

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

50

Gln Gly Ile Gly Leu Ala Ile Ala Asp Lys Phe Val Glu Glu Gly Ala
 20 25 30

55

EP 2 566 497 B1

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

<210> 147

<211> 759

55 <212> DNA

<213> Artificial Sequence

<220>

EP 2 566 497 B1

<223> Variant of L. kefir ketoreductase

<400> 147

```

5  atgaccaacc gtctgaagag caaagtagcc atcgtaaccg gcgggaccca gggatatcgg      60
   ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttatcac cggtcgccgt      120
   gcagatgtag gtgaaaaggc cgccaaatca atcggcggta ctgatgttat tcgctttgtc      180
10  cagcacgatg tgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca      240
   ttcgggcccg ttacgaccgt cgtgaacaat gcagggatct ctctgcctaa aagcgttgaa      300
   gacactacca cggaggaatg gcgcaaactg ttgtccgta atctggatag tgtttttttc      360
   ggcacccgtc tgggcattcg ccgcatgaag aataaaggct tgggcgctag catcatcaat      420
   atgagctcga tcagtgggct ggtaggcgat ccgatgttgg gggcatacaa cgcttccaag      480
20  ggggcgggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat      540
   gtgcggtgtca acacagtaca tccgggcgct atcaagacc cgctgactga taaatttcca      600
   gcaggtgggg aaatgatctc acagcgtacg aaaacccta tgggccacat tggngaaccg      660
25  aatgacgtgg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt      720
   agcgaatttg tggtcgacgg cgggtatacc gcacagtga      759

```

```

30  <210> 148
     <211> 252
     <212> PRT
     <213> Artificial Sequence

```

```

35  <220>
     <223> Variant of L. kefir ketoreductase

     <400> 148

```

40

45

50

55

EP 2 566 497 B1

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

EP 2 566 497 B1

<210> 149
 <211> 759
 <212> DNA
 <213> Artificial Sequence

5

<220>
 <223> Variant of L. kefir ketoreductase

<400> 149

10

```

atgaccaacc gtctgaagag caaagtagcc atcgtaaccg gcgggaccca gggatatcgg      60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttatcac cggtcgccgt      120
gcagatgtag gtgaaaaggc cgccaaatca atcggcggta ctgatgttat tcgctttgtc      180
cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca      240
ttcgggccgg ttacgaccgt cgtgaacaat gcagggatct ctctgcctaa aagcgttgaa      300
gacactacca cggaggaatg gcgcaaactg ttgtccgta atctggatag tgtttttttc      360
ggcaccocgtc tgggcattcg ccgcatgaag aataaaggct tgggcgctag catcatcaat      420
atgagctcga tcagtgggct ggtaggcgat ccgatgttgg gggcatacaa cgcttccaag      480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat      540
gtgcgtgtca acacagtaca tccgggcgct atcaagacc cgtgactga taaatttcca      600
gcaggtgggg aaatgatctc acagcgtacg aaaacccta tgggccacat tgggtgaaccg      660
aatgacattg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt      720
agcgaatttg tggtcgacgg cgggtatacc gcacagtga      759
  
```

35

<210> 150
 <211> 252
 <212> PRT
 <213> Artificial Sequence

40

<220>
 <223> Variant of L. kefir ketoreductase

<400> 150

45

50

55

EP 2 566 497 B1

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

EP 2 566 497 B1

<210> 151
 <211> 759
 <212> DNA
 <213> Artificial Sequence

5

<220>
 <223> Variant of L. kefir ketoreductase

<400> 151

10

atgaccaacc	gtctgaagag	caaagtagcc	atcgtaaccg	gcgggaccca	gggtatcggt	60
ttggcaatcg	ccgataaatt	tgtagaggag	ggtgcgaaag	tagttatcac	cggtcgccgt	120
gcagatgtag	gtgaaaaggc	cgccaaatca	atcggcggtg	ctgatgttat	tcgctttgtc	180
cagcacgatg	catccgatga	agcaggctgg	acgaaactgt	tcgacaccac	cgaggaggca	240
ttcggcccgg	ttacgaccgt	cgtgaacaat	gcagggatct	ctctgcctaa	aagcgttgaa	300
gacactacca	cggaggaatg	gcgcaaactg	ttgtccgtta	atctggatag	tgtttttttc	360
ggcaccgcgc	tgggcattcg	ccgcatgaag	aataaaggct	tgggcgctag	catcatcaat	420
atgagctcga	tcagtgggct	ggtaggcgat	ccgatgttgg	gggcatacaa	cgcttccaag	480
ggggcggtac	gtatcatgtc	gaaaagcgca	gcgctggatt	gcgcagtga	ggactacgat	540
gtgcgtgtca	acacagtaca	tccgggcgct	atcaagaccc	cgatgactga	taaatttcca	600
gcaggtgggg	aaatgatctc	acagcgtacg	aaaaccacct	tgggccacat	tggatgaaccg	660
aatgacgtgg	catggatctg	tgtgtacctg	gcacctgacg	aatcgaaatt	tgcgacgggt	720
agcgaatttg	tggtcgacgg	cgggtatacc	gcacagtga			759

35

<210> 152
 <211> 252
 <212> PRT
 <213> Artificial Sequence

40

<220>
 <223> Variant of L. kefir ketoreductase

<400> 152

45

Met Thr Asn Arg Leu Lys Ser Lys Val Ala Ile Val Thr Gly Gly Thr

50

55

EP 2 566 497 B1

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

EP 2 566 497 B1

<210> 153
 <211> 759
 <212> DNA
 <213> Artificial Sequence

5

<220>
 <223> Variant of L. kefir ketoreductase

<400> 153

10

atgaccaacc	gtctgaagag	caaagtagcc	atcgtaaccg	gcgggaccca	gggtatcggt	60
ttggcaatcg	ccgataaatt	tgtagaggag	ggtgcgaaag	tagttatcac	cggtcgcgct	120
gcagatgtag	gtgaaaaggc	cgccaaatca	atcggcggtg	ctgatgttat	tcgctttgtc	180
cagcacgatg	catccgatga	agcaggctgg	acgaaactgt	tcgacaccac	cgaggaggca	240
ttcggcccgg	ttacgaccgt	cgtgaacaat	gcagggatct	ctctgcctaa	aagcgttgaa	300
gacactacca	cggaggaatg	gcgcaaactg	ttgtccgtta	atctggatag	tgtttttttc	360
ggcaccgctc	tgggcattcg	ccgcatgaag	aataaaggct	tgggcgctag	catcatcaat	420
atgagctcga	tcagtgggat	tgtaggcgat	ccgatgttgg	gggcatacaa	cgcttccaag	480
ggggcggtac	gtatcatgtc	gaaaagcgca	gcgctggatt	gcgcagtga	ggactacgat	540
gtgcgtgtca	acacagtaca	tccgggcgct	atcaagaccc	cgctgactga	taaatttcca	600
gcaggtgggg	aaatgatctc	acagcgtacg	aaaaccctta	tgggccacat	tggatgaaccg	660
aatgacgtgg	catggatctg	tgtgtacctg	gcctctgacg	aatcgaaatt	tgcgacgggt	720
agcgaatttg	tggtcgacgg	cgggtatacc	gcacagtga			759

35

<210> 154
 <211> 252
 <212> PRT
 <213> Artificial Sequence

40

<220>
 <223> Variant of L. kefir ketoreductase

<400> 154

45

50

55

EP 2 566 497 B1

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

5 Gln Gly Ile Gly Leu Ala Ile Ala Asp Lys Phe Val Glu Glu Gly Ala
20 25 30

10 Lys Val Val Ile Thr Gly Arg Arg Ala Asp Val Gly Glu Lys Ala Ala
35 40 45

15 Lys Ser Ile Gly Gly Thr Asp Val Ile Arg Phe Val Gln His Asp Ala
50 55 60

Ser Asp Glu Ala Gly Trp Thr Lys Leu Phe Asp Thr Thr Glu Glu Ala
65 70 75 80

20 Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ile Ser Leu Pro

25

30

35

40

45

50

55

EP 2 566 497 B1

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

<210> 155

<211> 759

<212> DNA

<213> Artificial Sequence

<220>

<223> Variant of L. kefir ketoreductase

<400> 155

EP 2 566 497 B1

	atgaccaacc gtctgaagag caaagtagcc atcgtaaccg gcgggaccca gggatatcgg	60
	ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttatcac cggtcgccgt	120
5	gcagatgtag gtgaaaaggc cgccaaatca atcggcggtta ctgatgttat tcgctttgtc	180
	cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca	240
	ttcggccccg ttacgaccgt cgtgaacaat gcagggatct ctctgcctaa aagcctggaa	300
10	gacactacca cggaggaatg gcataaactg ttgtccgtta atctggatag tgtttttttc	360
	ggcaccogtc tgggcattcg ccgcatgaag aataaaggct tgggcgctag catcatcaat	420
15	atgagctoga tcagtgggat cgtaggcgat ccgatgttg gggcatacaa cgcttccaag	480
	ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat	540
	gtgcgtgtca acacagtaca tccgggcgct atcaagacc cgctgactga taaatttcca	600
20	gcaggtgggg aagtgatctc acagcgtagc aaaacccta tgggccacat tggngaaccg	660
	aatgacgtgg catgggtgtg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt	720
25	agcgaatttg tggtcgacgg cgggtatacc gcacagtga	759

<210> 156

<211> 252

<212> PRT

30 <213> Artificial Sequence

<220>

<223> Variant of L. kefir ketoreductase

35 <400> 156

40

45

50

55

EP 2 566 497 B1

	Met	Thr	Asn	Arg	Leu	Lys	Ser	Lys	Val	Ala	Ile	Val	Thr	Gly	Gly	Thr
	1				5					10					15	
5	Gln	Gly	Ile	Gly	Leu	Ala	Ile	Ala	Asp	Lys	Phe	Val	Glu	Glu	Gly	Ala
				20					25					30		
10	Lys	Val	Val	Ile	Thr	Gly	Arg	Arg	Ala	Asp	Val	Gly	Glu	Lys	Ala	Ala
			35					40					45			
15	Lys	Ser	Ile	Gly	Gly	Thr	Asp	Val	Ile	Arg	Phe	Val	Gln	His	Asp	Ala
		50					55					60				
20	Ser	Asp	Glu	Ala	Gly	Trp	Thr	Lys	Leu	Phe	Asp	Thr	Thr	Glu	Glu	Ala
	65					70					75					80
25	Phe	Gly	Pro	Val	Thr	Thr	Val	Val	Asn	Asn	Ala	Gly	Ile	Ser	Leu	Pro
					85					90					95	
30	Lys	Ser	Leu	Glu	Asp	Thr	Thr	Thr	Glu	Glu	Trp	His	Lys	Leu	Leu	Ser
				100					105					110		
35	Val	Asn	Leu	Asp	Ser	Val	Phe	Phe	Gly	Thr	Arg	Leu	Gly	Ile	Arg	Arg
			115					120					125			
40	Met	Lys	Asn	Lys	Gly	Leu	Gly	Ala	Ser	Ile	Ile	Asn	Met	Ser	Ser	Ile
		130					135					140				
45	Ser	Gly	Ile	Val	Gly	Asp	Pro	Met	Leu	Gly	Ala	Tyr	Asn	Ala	Ser	Lys
	145					150					155					160
50	Gly	Ala	Val	Arg	Ile	Met	Ser	Lys	Ser	Ala	Ala	Leu	Asp	Cys	Ala	Val
55																

EP 2 566 497 B1

	165	170	175
5	Lys Asp Tyr Asp Val Arg Val Asn Thr Val His Pro Gly Ala Ile Lys 180 185 190		
10	Thr Pro Leu Thr Asp Lys Phe Pro Ala Gly Gly Glu Val Ile Ser Gln 195 200 205		
15	Arg Thr Lys Thr Pro Met Gly His Ile Gly Glu Pro Asn Asp Val Ala 210 215 220		
20	Trp Val Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly 225 230 235 240		
25	Ser Glu Phe Val Val Asp Gly Gly Tyr Thr Ala Gln 245 250		
30	<210> 157 <211> 759 <212> DNA <213> Artificial Sequence <220> <223> Variant of L. kefir ketoreductase <400> 157		
35	atgaccaacc gtctgaagag caaagtagcc atcgtaaccg gcgggaccca gggatatcggg	60	
40	ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttatcac cggtcgccgt	120	
45	gcagatgtag gtgaaaaggc cgccaaatca atcggcggta ctgatgttat tcgctttgtc	180	
50	cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca	240	
55	ttcggcccgg ttacgaccgt cgtgaacaat gcagggatct ctctgcctaa aagcctggaa	300	
	gacactacca cggaggaatg gcataaactg ttgtccgtta atctggatag tgtttttttc	360	
	ggcaccgcgc tgggcattcg ccgcatgaag aataaaggct tgggcgctag catcatcaat	420	
	atgagctcga tcagtgggat cgtaggcgat ccgatgttgg gggcatacaa cgcttccaag	480	
	ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat	540	
	gtgcgtgtca acacagtaca tccgggcgct atcaagaccc cgctgactga taaatttcca	600	
	gcaggtgggg aaatgatctc acagcgtacg aaaacccta tgggccacat tgggtgaaccg	660	
	aatgacgtgg catgggtgtg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt	720	
	agcgaatttg tggtcgacgg cgggtatacc gcacagtga	759	
	<210> 158 <211> 252		

EP 2 566 497 B1

<212> PRT

<213> Artificial Sequence

<220>

5 <223> Variant of L. kefir ketoreductase

<400> 158

10

15

20

25

30

35

40

45

50

55

EP 2 566 497 B1

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

245

250

5 <210> 159
 <211> 759
 <212> DNA
 <213> Artificial Sequence

 10 <220>
 <223> Variant of L. kefir ketoreductase

 <400> 159

 15 atgaccaacc gtctgaagag caaagtagcc atcgtaaccg gcgggaccca gggatatcgg 60
 ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttatcac cggtcgcccg 120
 gcagatgtag gtgaaaaggc cgccaaatca atcggcggta ctgatgttat tcgctttgtc 180
 20 cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
 ttccggcccg ttacgaccgt cgtgaacaat gcagggatct ctctgcctaa aagcctggaa 300
 gacactacca cggaggaatg gcgcaaactg ttgtccgta atctggatag tgtttttttc 360
 25 ggcacccgtc tgggcattcg ccgcatgaag aataaaggct tgggcgctag catcatcaat 420
 atgagctcga tcagtgggat cgtaggcgat ccgatgttgg gggcatacaa cgcttccaag 480
 30 ggggcgggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat 540
 gtgcgtgtca acacagtaca tccgggcgct atcaagacc cgctgactga taaatttcca 600
 gcagggtggg aagtgatctc acagcgtacg aaaacccta tgggocacat tggatgaaccg 660
 35 aatgacgtgg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
 agcgaatttg tggtcgacgg cgggtatacc gcacagtga 759

40 <210> 160
 <211> 252
 <212> PRT
 <213> Artificial Sequence

45 <220>
 <223> Variant of L. kefir ketoreductase

 <400> 160

50

55

EP 2 566 497 B1

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

EP 2 566 497 B1

<210> 161
 <211> 759
 <212> DNA
 <213> Artificial Sequence

5

<220>
 <223> Variant of L. kefir ketoreductase

<400> 161

10

atgaccaacc	gtctgaagag	caaagtagcc	atcgtaaccg	gcgggaccca	gggtatcgg	60
ttggcaatcg	ccgataaatt	tgtagaggag	ggtgcgaaag	tagttatcac	cggtcgccgt	120
gcagatgtag	gtgaaaaggc	cgccaaatca	atcggcggtg	ctgatgttat	tcgctttgtc	180
cagcacgatg	catccgatga	agcaggctgg	acgaaactgt	tcgacaccac	cgaggaggca	240
ttcggcccg	ttacgaccgt	cgtgaacaat	gcagggatct	ctctgcctaa	aagcctggaa	300
gacactacca	cggaggaatg	gcataaactg	ttgtccgtta	atctggatgg	tgtttttttc	360
ggcacccgtc	tgggcattcg	ccgcatgaag	aataaaggct	tgggcgctag	catcatcaat	420
atgagctcga	tcagtgggat	cgtaggcgat	ccgatgttgg	gggcatacaa	cgcttccaag	480
ggggcggtac	gtatcatgtc	gaaaagcgca	gcgctggatt	gcgcagtga	ggactacgat	540
gtgcgtgtca	acacagtaca	tccgggcgct	atcaagaccc	cgctgactga	taaattttcca	600
gcagggtggg	aaatgatctc	acagcgtacg	aaaaccctta	tgggccacat	tggtgaaccg	660
aatgacgtgg	catgggtgtg	tgtgtacctg	gcattctgacg	aatcgaaatt	tgcgacgggt	720
agcgaatttg	tggtcgacgg	cgggtatacc	gcacagtga			759

35

<210> 162
 <211> 252
 <212> PRT
 <213> Artificial Sequence

40

<220>
 <223> Variant of L. kefir ketoreductase

<400> 162

45

50

55

EP 2 566 497 B1

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

EP 2 566 497 B1

<210> 163
 <211> 759
 <212> DNA
 <213> Artificial Sequence

5

<220>
 <223> Variant of L. kefir ketoreductase

<400> 163

10

```

atgaccaacc gtctgaagag caaagtagcc atcgtaaccg gcgggaccca gggatatcgg      60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttatcac cggtcgccgt      120
gcagatgtag gtgaaaaggc cgccaaatca atcggcggta ctgatgttat tcgctttgtc      180
cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca      240
ttcggccccg ttacgaccgt cgtgaacaat gcagggatct ctctgcctaa aagcctggaa      300
gacactacca cggaggaatg gcataaactg ttgtccgta atctggatag tgtttttttc      360
ggcaccctgc tgggcattcg ccgcatgaag aataaaggct tgggcgctag catcatcaat      420
atgagctcga tcagtgggat cgtaggcgat ccgatgttgg gggcatacaa cgcttccaag      480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat      540
gtgcgtgtca acacagtaca tccgggcgct atcaagacc cgctgactga taaatttcca      600
gcaggtgggg aagtgatctc acagcgtagc aaaaccctta tgggccacat tggggaaccg      660
aatgacgtgg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt      720
agcgaatttg tggtcgacgg cgggtatacc gcacagtga      759

```

35

<210> 164
 <211> 252
 <212> PRT
 <213> Artificial Sequence

40

<220>
 <223> Variant of L. kefir ketoreductase

<400> 164

45

50

55

EP 2 566 497 B1

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

EP 2 566 497 B1

<210> 165
 <211> 759
 <212> DNA
 <213> Artificial Sequence

5

<220>
 <223> Variant of L. kefir ketoreductase

<400> 165

10

```

atgaccaacc gtctgaagag caaagtagcc atcgtaaccc gcgggaccca gggatcgggt      60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttatcac cggtcgccgt      120
gcagatgtag gtgaaaaggg cgccaaatca atcggcggtta ctgatgttat tcgctttgtc      180
cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca      240
ttcggcccggt ttacgaccgt cgtgaacaat gcagggatct ctctgcctaa aagcctggaa      300
gacactacca cggaggaatg gcgcaaactg ttgtccgtta atctggatag tgtttttttc      360
ggcaccgcgtc tgggcattcg ccgcatgaag aataaaggct tgggcgctag catcatcaat      420
atgagctcga tcagtgggat cgtaggcgat ccgatgttgg gggcatacaa cgcttccaag      480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat      540
gtgcgtgtca acacagtaca tccgggcgct atcaagacc cgctgactga taaatttcca      600
gcaggtgggg aagtgatctc acagcgtacg aaaaccctta tgggccacat tggatgaaccg      660
aatgacgtgg catgggtgtg tgtgtacctg gcctctgacg aatcgaaatt tgcgacgggt      720
agcgaatttg tggtcgacgg cgggtatacc gcacagtga      759
  
```

35

<210> 166
 <211> 252
 <212> PRT
 <213> Artificial Sequence

40

<220>
 <223> Variant of L. kefir ketoreductase

<400> 166

45

```

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

Gln Gly Ile Gly Leu Ala Ile Ala Asp Lys Phe Val Glu Glu Gly Ala
20           25           30

Lys Val Val Ile Thr Gly Arg Arg Ala Asp Val Gly Glu Lys Ala Ala
35           40           45
  
```

55

EP 2 566 497 B1

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

<210> 167

<211> 759

<212> DNA

<213> Artificial Sequence

<220>

<223> Variant of L. kefir ketoreductase

<400> 167

atgaccaacc gtctgaagag caaagtagcc atcgtaaccg gcgggaccca gggatatcgg 60
 ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttatcac cggtcgccgt 120
 5 gcagatgtag gtgaaaaggc cgccaaatca atcggcggtta ctgatgttat tcgctttgtc 180
 cagcacgatg tgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
 10 ttccggcccg ttacgaccgt cgtgaacaat gcagggatct ctctgcctaa aagcctggaa 300
 gacactacca cggaggaatg gcataaactg ttgtccgtta atctggatgg tgtttttttc 360
 ggcaccgcgc tgggcattcg ccgcatgaag aataaaggct tgggcgctag catcatcaat 420
 15 atgagctcga tcagtgggat cgtaggcgat ccgatgttgg gggcatacaa cgcttccaag 480
 ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcagtga ggactacgat 540
 20 gtgcgtgtca acacagtaca tccgggcgct atcaagacct cgctgactga taaatggcca 600
 gcaggtgggg aaatgatctc acagcgtagc aaaacccta tgggccacat tggatgaaccg 660
 aatgacgtgg catgggtgtg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
 25 agcgaatttg tggtcgacgg cgggtatacc gcacagtga 759

<210> 168

<211> 252

<212> PRT

30 <213> Artificial Sequence

<220>

<223> Variant of L. kefir ketoreductase

35 <400> 168

40

45

50

55

EP 2 566 497 B1

	Met	Thr	Asn	Arg	Leu	Lys	Ser	Lys	Val	Ala	Ile	Val	Thr	Gly	Gly	Thr
	1				5					10					15	
5	Gln	Gly	Ile	Gly	Leu	Ala	Ile	Ala	Asp	Lys	Phe	Val	Glu	Glu	Gly	Ala
				20					25					30		
10	Lys	Val	Val	Ile	Thr	Gly	Arg	Arg	Ala	Asp	Val	Gly	Glu	Lys	Ala	Ala
			35					40					45			
15	Lys	Ser	Ile	Gly	Gly	Thr	Asp	Val	Ile	Arg	Phe	Val	Gln	His	Asp	Val
		50					55					60				
20	Ser	Asp	Glu	Ala	Gly	Trp	Thr	Lys	Leu	Phe	Asp	Thr	Thr	Glu	Glu	Ala
	65					70					75					80
25	Phe	Gly	Pro	Val	Thr	Thr	Val	Val	Asn	Asn	Ala	Gly	Ile	Ser	Leu	Pro
					85					90					95	
30	Lys	Ser	Leu	Glu	Asp	Thr	Thr	Thr	Glu	Glu	Trp	His	Lys	Leu	Leu	Ser
				100					105					110		
35	Val	Asn	Leu	Asp	Gly	Val	Phe	Phe	Gly	Thr	Arg	Leu	Gly	Ile	Arg	Arg
			115					120					125			

Met Lys Asn Lys Gly Leu Gly Ala Ser Ile Ile Asn Met Ser Ser Ile
 130 135 140

Ser Gly Ile Val Gly Asp Pro Met Leu Gly Ala Tyr Asn Ala Ser Lys
 145 150 155 160

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

Lys Asp Tyr Asp Val Arg Val Asn Thr Val His Pro Gly Ala Ile Lys
 180 185 190

Thr Pro Leu Thr Asp Lys Trp Pro Ala Gly Gly Glu Met Ile Ser Gln
 195 200 205

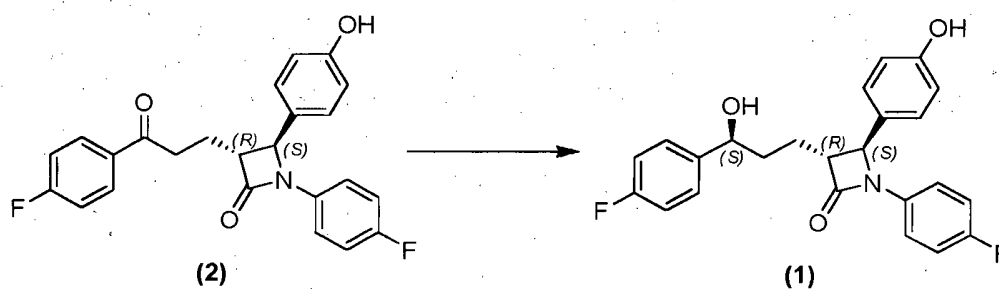
Arg Thr Lys Thr Pro Met Gly His Ile Gly Glu Pro Asn Asp Val Ala
 210 215 220

Trp Val Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly
 225 230 235 240

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

Claims

1. A non-naturally occurring polypeptide capable of converting the ketone of compound (2) to the chiral alcohol of compound (1) in diastereomeric excess



the polypeptide comprising an amino acid sequence that has at least 80% sequence identity to SEQ ID NO:2 and at least the following features:

residue at position corresponding to X206 is I;
 residue at position corresponding to X40 is R;
 residue at position corresponding to X153 is I, or L;
 residue at position corresponding to X190 is A or P;
 residue at position corresponding to X196 is T; and
 residue at position corresponding to X199 is F, or W.

2. The polypeptide of claim 1 in which:

(i) the amino acid sequence further comprises a group of features selected from:

5 residue at position X93 is A and at position X94 is T;
 residue at position X93 is A and at position X94 is S;
 residue at position X93 is A and at position X94 is P; or
 residue at position X93 is I and at position X94 is S;

10 (ii) the amino acid sequence comprises at least the following features:

 residue at position corresponding to X40 is R;
 residue at position corresponding to X93 is A;
 residue at position corresponding to X94 is T;
 15 residue at position corresponding to X153 is I;
 residue at position corresponding to X190 is A;
 residue at position corresponding to X196 is T;
 residue at position corresponding to X199 is F;
 residue at position corresponding to X203 is G; and
 20 residue at position corresponding to X206 is I;

(iii) the amino acid sequence comprises the following features:

 residue at position corresponding to X40 is R;
 25 residue at position corresponding to X93 is A;
 residue at position corresponding to X94 is S;
 residue at position corresponding to X96 is N;
 residue at position corresponding to X153 is L;
 residue at position corresponding to X190 is A;
 30 residue at position corresponding to X196 is T;
 residue at position corresponding to X199 is F;
 residue at position corresponding to X202 is G;
 residue at position corresponding to X203 is G; and
 residue at position corresponding to X206 is I;

35 (iv) the amino acid sequence comprises the following features:

 residue at position corresponding to X40 is R;
 residue at position corresponding to X93 is A;
 40 residue at position corresponding to X94 is S;
 residue at position corresponding to X96 is A, P, or N;
 residue at position corresponding to X153 is L;
 residue at position corresponding to X190 is A;
 residue at position corresponding to X196 is T;
 45 residue at position corresponding to X199 is F;
 residue at position corresponding to X202 is G;
 residue at position corresponding to X203 is G; and residue at position corresponding to X206 is I,
 optionally in which the amino acid sequence further comprises the following features:

50 residue at position corresponding to X96 is N; and
 residue at position corresponding to X201 is I, or L;

(v) the amino acid sequence comprises at least the following features:

55 residue at position corresponding to X40 is R;
 residue at position corresponding to X93 is A;
 residue at position corresponding to X96 is P;
 residue at position corresponding to X153 is L;

residue at position corresponding to X190 is A;
 residue at position corresponding to X196 is T;
 residue at position corresponding to X199 is F;
 residue at position corresponding to X202 is G;
 residue at position corresponding to X203 is G; and
 residue at position corresponding to X206 is I,
 optionally in which the amino acid sequence further comprises one or more of the following features:

residue at position corresponding to X21 is F;
 residue at position corresponding to X25 is R;
 residue at position corresponding to X94 is S;
 residue at position corresponding to X96 is P;
 residue at position corresponding to X108 is K;
 residue at position corresponding to X127 is K;
 residue at position corresponding to X147 is M;
 residue at position corresponding to X153 is L;
 residue at position corresponding to X205 is V;
 residue at position corresponding to X211 is K;
 residue at position corresponding to X221 is D;

(vi) the amino acid sequence comprises at least the following features:

residue at position corresponding to X21 is F;
 residue at position corresponding to X25 is R;
 residue at position corresponding to X40 is R;
 residue at position corresponding to X93 is A;
 residue at position corresponding to X94 is S;
 residue at position corresponding to X96 is P;
 residue at position corresponding to X108 is K;
 residue at position corresponding to X127 is K;
 residue at position corresponding to X147 is M;
 residue at position corresponding to X153 is L;
 residue at position corresponding to X190 is A;
 residue at position corresponding to X196 is T;
 residue at position corresponding to X199 is F;
 residue at position corresponding to X202 is G;
 residue at position corresponding to X203 is G;
 residue at position corresponding to X206 is I; and
 residue at position corresponding to X221 is D;

(vii) the amino acid sequence comprises at least the following features:

residue at position corresponding to X40 is R;
 residue at position corresponding to X93 is I;
 residue at position corresponding to X94 is S;
 residue at position corresponding to X96 is P;
 residue at position corresponding to X153 is L;
 residue at position corresponding to X190 is A;
 residue at position corresponding to X196 is T;
 residue at position corresponding to X199 is F, or W;
 residue at position corresponding to X201 is A;
 residue at position corresponding to X202 is G;
 residue at position corresponding to X203 is G;
 residue at position corresponding to X206 is I; and residue at position corresponding to X211 is K,
 optionally in which the amino acid sequence further comprises one or more of the following features:

residue at position corresponding to X64 is V;
 residue at position corresponding to X99 is L;

residue at position corresponding to X108 is H;
 residue at position corresponding to X117 is A, or G;
 residue at position corresponding to X147 is I, or M;
 residue at position corresponding to X195 is M;
 residue at position corresponding to X205, is V;
 residue at position corresponding to X223 is I;
 residue at position corresponding to X226 is V;

(viii) the amino acid sequence comprises at least the following features:

residue at position corresponding to X40 is R;
 residue at position corresponding to X93 is I;
 residue at position corresponding to X94 is S;
 residue at position corresponding to X96 is P;
 residue at position corresponding to X99 is L;
 residue at position corresponding to X108 is H;
 residue at position corresponding to X117 is G;
 residue at position corresponding to X147 is I;
 residue at position corresponding to X153 is L;
 residue at position corresponding to X190 is A;
 residue at position corresponding to X196 is T;
 residue at position corresponding to X199 is F;
 residue at position corresponding to X201 is A;
 residue at position corresponding to X202 is G;
 residue at position corresponding to X203 is G;
 residue at position corresponding to X206 is I;
 residue at position corresponding to X211 is K; and
 residue at position corresponding to X226 is V; or

(ix) the amino acid sequence comprises the following features:

residue at position corresponding to X40 is R;
 residue at position corresponding to X64 is V;
 residue at position corresponding to X93 is I;
 residue at position corresponding to X94 is S;
 residue at position corresponding to X96 is P;
 residue at position corresponding to X99 is L;
 residue at position corresponding to X108 is H;
 residue at position corresponding to X117 is G;
 residue at position corresponding to X147 is I;
 residue at position corresponding to X153 is L;
 residue at position corresponding to X190 is A;
 residue at position corresponding to X196 is T;
 residue at position corresponding to X199 is W;
 residue at position corresponding to X201 is A;
 residue at position corresponding to X202 is G;
 residue at position corresponding to X203 is G;
 residue at position corresponding to X206 is I;
 residue at position corresponding to X211 is K; and
 residue at position corresponding to X226 is V.

3. The polypeptide of claim 1 or 2(i) in which the amino acid sequence further comprises a feature or group of features selected from:

residue at position X203 is G;
 residue at position X202 is G and at position X203 is G; or
 residue at position X201 is A, position X202 is G and at position X203 is G.

4. The polypeptide of any one of claims 1 to 3 in which the amino acid sequence further comprises one or more of the following features:

residue at position corresponding to X21 is R or F;
 residue at position corresponding to X25 is R, T, or N;
 residue at position corresponding to X40 is R;
 residue at position corresponding to X64 is V;
 residue at position corresponding to X93 is A;
 residue at position corresponding to X94 is T, S, or P;
 residue at position corresponding to X95 is V, or M;
 residue at position corresponding to X96 is V, G, A, N, S, P, or T;
 residue at position corresponding to X99 is L;
 residue at position corresponding to X108 is H;
 residue at position corresponding to X117 is A, or G;
 residue at position corresponding to X127 is K, or Q;
 residue at position corresponding to X147 is M, or I;
 residue at position corresponding to X148 is I;
 residue at position corresponding to X150 is H, or A;
 residue at position corresponding to X152 is N, or F;
 residue at position corresponding to X153 is I, or L;
 residue at position corresponding to X155 is C;
 residue at position corresponding to X190 is A;
 residue at position corresponding to X195 is M;
 residue at position corresponding to X196 is T, A, S, C, or N;
 residue at position corresponding to X199 is F, or W;
 residue at position corresponding to X201 is I, L, or A;
 residue at position corresponding to X202 is L, N, V, or G;
 residue at position corresponding to X203 is G;
 residue at position corresponding to X204 is V, or A;
 residue at position corresponding to X205 is V;
 residue at position corresponding to X206 is I;
 residue at position corresponding to X207 is T, C, I, or N;
 residue at position corresponding to X211 is K;
 residue at position corresponding to X221 is D;
 residue at position corresponding to X223 is I;
 residue at position corresponding to X226 is V.

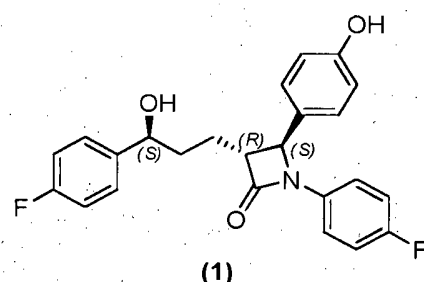
5. The polypeptide of claim 1 in which:

(i) the polypeptide is capable of converting compound (2) to compound (1) in at least 99% diastereomeric excess;
 (ii) the polypeptide is capable of at least about 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or greater conversion of compound (2) to compound (1) in 24 h with a substrate loading of about 100 g/L;
 (iii) the polypeptide is capable of converting compound (2) to compound (1) with at least 2-fold, at least 10-fold, at least 25-fold, at least 40-fold, or at least 60-fold increased activity relative to the activity of the polypeptide of SEQ ID NO: 2;
 (iv) the polypeptide is capable of converting compound (2) to compound (1) in a diastereomeric excess of compound (1) of at least 97%, 98, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, 99.6%, 99.7%, 99.8%, or 99.9% or greater in 24 h with a substrate loading of about 100 g/L, optionally in which the reaction conditions further comprise an IPA concentration of about 55-75% (v/v), an NADPH or NADH cofactor loading of about 0.03-0.5 g/L, and wherein no cofactor recycling enzyme present other than the non-naturally occurring polypeptide; or
 (v) the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO: 168, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164 and 166.

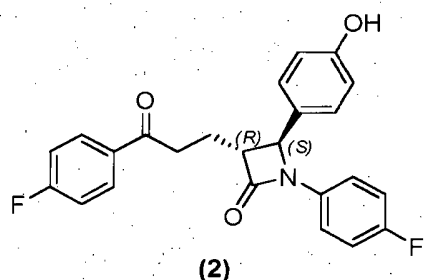
6. A polynucleotide encoding the polypeptide of any one of claims 1 to 5, optionally which comprises a nucleotide sequence selected from the group consisting of SEQ ID NO: 167, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29,

31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139, 141, 143, 145, 147, 149, 151, 153, 155, 157, 159, 161, 163, or 165.

7. An expression vector comprising the polynucleotide of claim 6, optionally comprising a control sequence, for example in which the control sequence comprises a promoter.
8. A host cell comprising a polynucleotide of claim 6 or an expression vector of claim 7, optionally which is *E. coli*.
9. A method for preparing a non-naturally occurring polypeptide of any one of claims 1 to 5 comprising culturing a host cell of claim 8 and isolating the non-naturally occurring polypeptide from the cell.
10. A method for preparing compound (1) or an analog of compound (1) in diastereomeric excess



comprising contacting compound (2) or an analog of compound (2)



with a polypeptide of any one of claims 1 to 5 in the presence of cofactor NADPH or NADH under suitable reaction conditions.

11. The method of claim 10 in which:

(i) the suitable reaction conditions comprise:

- (1) substrate loading of about 50-200 g/L compound (2);
- (2) polypeptide concentration of about 1.5-2.5 g/L;
- (3) NADPH cofactor concentration of about 0.1-0.2 g/L;
- (4) a co-solvent solution of an aqueous buffer and about 60-70% (v/v) IPA;
- (5) about pH 6.6-7.0; and
- (6) temperature of about 28-30 °C;

- (ii) the suitable reaction conditions comprise a compound (2) loading of at least about 20 g/L, about 40 g/L, about 50 g/L, about 75 g/L, about 100 g/L, about 200 g/L, about 250 g/L, about 300 g/L, or about 400 g/L;
- (iii) the suitable reaction conditions comprise an engineered polypeptide concentration of about 0.1-3.0 g/L, about 0.5-2.75 g/L, about 1.0-2.5 g/L, about 1.5-2.5 g/L, about 3 g/L, about 2 g/L, about 1.5 g/L, about 1.0 g/L, about 0.75 g/L, or even lower concentration;
- (iv) the suitable reaction conditions comprise a pH of about 6.0 to about 7.5, a pH of about 6.25 to about 7.25,

a pH of about 6.5 to about 7.25, a pH of about 6.6 to about 7.25, a pH of about 6.6 to about 7.0, a pH of about 6.75 to about 7.25, or a pH of about 6.75;

(v) the suitable reaction conditions comprise a temperature of about 20°C to about 40°C, about 23°C to about 37°C, about 25°C to about 35°C, about 26°C to about 32°C, or about 28°C to about 30°C;

(vi) the suitable reaction conditions comprise an aqueous buffer solution, an organic solvent, or a co-solvent system;

(vii) the suitable reaction conditions comprise a buffer solution of about 0.05 M TEA to about 0.25 M TEA;

(viii) the suitable reaction conditions comprise a co-solvent system comprising about 30% (v/v) to about 70% (v/v) of an aqueous buffer solution and about 30% (v/v) to about 70% (v/v) of an organic solvent solution, optionally in which the co-solvent system comprises an aqueous buffer solution, IPA, and another organic solvent, such as in which:

(a) the co-solvent system comprises about 45-55% (v/v) of an aqueous buffer solution, about 15-25% (v/v) IPA, and about 25-35% (v/v) toluene; or

(b) the co-solvent system comprises about 50% (v/v) of an aqueous buffer solution, about 20% (v/v) IPA, and about 30% (v/v) toluene;

(ix) the suitable reaction conditions comprise a co-solvent system comprising an aqueous buffer solution and IPA, wherein the IPA concentration is about 25-75% (v/v), about 35-75% (v/v), about 45-75% (v/v), about 55-75% (v/v), about 60-70% (v/v), about 62-68% (v/v), optionally in which the co-solvent system comprises 0.1 M TEA buffer and about 60% (v/v) to about 70% (v/v) IPA, for example in which the co-solvent system comprises about 35% (v/v) 0.1 M TEA buffer and about 65% (v/v) IPA;

(x) the suitable reaction conditions comprise an NADH or NADPH cofactor concentration of about 0.03-0.5 g/L, about 0.05-0.3 g/L, about 0.1-0.2 g/L, about 0.5 g/L, about 0.1 g/L, or about 0.2 g/L;

(xi) the method further comprises converting NADP⁺ to NADPH with a cofactor recycling system, optionally in which the cofactor recycling system is selected from a ketoreductase and a secondary alcohol, or glucose dehydrogenase and glucose, for example in which:

(a) the secondary alcohol is IPA; or

(b) the ketoreductase is the polypeptide of any one of claims 1 to 5;

(xii) the suitable reaction conditions comprise an IPA concentration is about 55-75% (v/v), an NADPH or NADH cofactor loading of about 0.03-0.5 g/L, and wherein no cofactor recycling enzyme is present other than the polypeptide of any one of claims 1 to 5;

(xiii) the suitable reaction conditions comprise:

(1) substrate loading of about 100 g/L compound (2);

(2) polypeptide concentration of about 2.0 g/L;

(3) NADPH cofactor concentration of about 0.1 g/L;

(4) a co-solvent solution of an aqueous buffer of 0.1 M TEA and about 65% v/v IPA;

(5) about pH 6.75; and

(6) temperature of about 30 °C;

(xiv) the suitable reaction conditions comprise:

(1) substrate loading of about 50-150 g/L of compound (2);

(2) polypeptide concentration of about 2.5-3.5 g/L;

(3) NADPH cofactor concentration of about 0.1-0.2 g/L;

(4) a co-solvent solution of an aqueous buffer and about 15-25% (v/v) IPA and 25-35% (v/v) toluene; .

(5) about a pH of 6.6-7.0; and

(6) a temperature of about 28-30 °C;

(xv) the suitable reaction conditions comprise:

(1) substrate loading of about 100 g/L of compound (2);

(2) polypeptide concentration of about 3.0 g/L;

(3) NADPH cofactor concentration of about 0.1 g/L;

(4) a co-solvent solution of an aqueous buffer of 0.1 M TEA and about 20% (v/v) IPA and 30% (v/v) toluene;

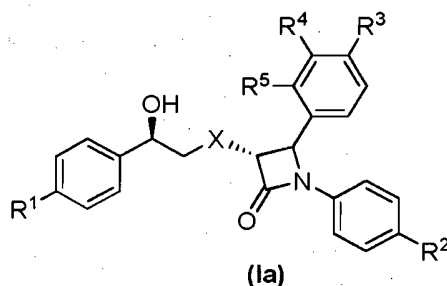
(5) about a pH of 6.6-7.0; and

(6) a temperature of about 28-30 °C;

(xvi) the suitable reaction conditions comprise a substrate loading of about 100 g/L compound (2) and wherein the method results in at least about 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or greater conversion of compound (2) to compound (1) in 24 h; or

(xvii) the suitable reaction conditions comprise a substrate loading of about 100 g/L compound (2) and wherein the method results in an diastereomeric excess of compound (1) of at least 97%, 98%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, 99.6%, 99.7%, 99.8%, or 99.9% or greater in 24 h.

12. The method claim 10 in which the analog of compound (1) is a compound of Formula Ia:



wherein,

X is CorS;

R¹ is selected from -H, -F, -Cl, -Br, or -I;

R² is selected from -H, -F, -Cl, -Br, -I, -CN, -OH (optionally protected with a hydroxyl protecting group), -CH₂NH₂ (optionally protected with a nitrogen protecting group), and any one of the following optionally substituted groups: alkyl, alkoxy, alkenyl, alkenoxy, alkynyl, alkynoxy, cycloalkyl, aryl, heteroaryl, or heterocycle;

R³ is selected from -H, -F, -Cl, -Br, -I, -CN, -OH (optionally protected with a hydroxyl protecting group), -CH₂NH₂ (optionally protected with a nitrogen protecting group), and any one of the following optionally substituted groups: alkyl, alkoxy, alkenyl, alkenoxy, alkynyl, alkynoxy, cycloalkyl, aryl, heteroaryl, or heterocycle;

R⁴ is selected from -H, -F, -Cl, -Br, -I, -CN, -OH (optionally protected with a hydroxyl protecting group); and

R⁵ is selected from -H, -F, -Cl, -Br, -I, -CN, -OH (optionally protected with a hydroxyl protecting group), -CH₂NH₂ (optionally protected with a nitrogen protecting group), and any one of the following optionally substituted groups: alkyl, alkoxy, alkenyl, alkenoxy, alkynyl, alkynoxy, cycloalkyl, aryl, heteroaryl, or heterocycle.

13. The method of claim 12 in which:

(i) R² and R³ are each independently selected from -H, -F, -Cl, -Br, -I, -CN, -OH (optionally protected with a hydroxyl protecting group), -CH₂NH₂ (optionally protected with a nitrogen protecting group), and any one of the following optionally substituted groups: -(C₁-C₆)alkyl, -(C₁-C₆)alkoxy, -(C₁-C₆)alkenyl, -(C₁-C₆)alkenoxy, -(C₁-C₆)alkynyl, -(C₁-C₆)alkynoxy, -(C₁-C₆)cycloalkyl, or a heterocycle, having from 1 to 4 carbon atoms and 1 to 2 hetero atoms;

(ii) R² and R³ are each independently selected from -H, -F, -Cl, -Br, -I, -CN, -OH (optionally protected with a hydroxyl protecting group), -CH₂NH₂ (optionally protected with a nitrogen protecting group), and any one of the following optionally substituted groups: -(C₁-C₄)alkyl, -(C₁-C₄)alkoxy, -(C₁-C₄)alkenyl, -(C₁-C₄)alkenoxy, -(C₁-C₄)alkynyl, -(C₁-C₄)alkynoxy, -(C₁-C₄)cycloalkyl, or a heterocycle, having from 1 to 3 carbon atoms and 1 to 2 hetero atoms;

(iii) R² and R³ are each independently selected from -H, -F, -Cl, -Br, -I, -CN, -OH (optionally protected with a hydroxyl protecting group), and -CH₂NH₂ (optionally protected with a nitrogen protecting group);

(iv) R² and R³ are each independently selected from the following optionally substituted groups: -(C₁-C₄)alkyl, -(C₁-C₄)alkoxy, -(C₁-C₄)alkenyl, -(C₁-C₄)alkenoxy, -(C₁-C₄)alkynyl, and -(C₁-C₄)alkynoxy;

(v) R² and R³ optionally substituted with one or more -OH groups; or

(vi) any one or more of R², R³, or R⁵ is an -OH or an -OH group protected with a hydroxyl protecting group, optionally wherein the hydroxyl protecting group is selected from the group consisting of benzyl, acetyl, benzoyl,

tert-butyloxycarbonyl, silyl, tert-butyldiphenylsilyl, trimethylsilyl, para-methoxybenzyl, benzylidene, dimethylacetal, and methoxy methyl, for example wherein the silyl group is -Si-(R^a)(R^b)(R^c) and R^a, R^b, and R^c are each independently selected from the group consisting of C₁-C₆ alkyl, phenyl, acetyl, and benzyl groups.

14. The method of claim 12 or claim 13 in which:

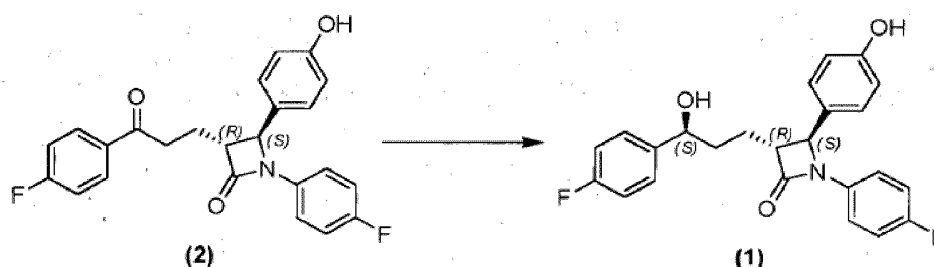
- (i) R¹ is -F, R⁴ is -H, and R⁵ is -H;
- (ii) R² is -F; and/or
- (iii) X is C or S.

15. The method of any one of claims 10 to 14 in which:

- (i) the method further comprises one or more steps selected from: extraction of compound (1); isolation of compound (1); purification of compound (1); and crystallization of compound (1); and/or
- (ii) the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO: 168, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, and 166.

Patentansprüche

1. Nichtnatürlich vorkommendes Polypeptid, das fähig ist, das Keton der Verbindung (2) mit Diastereomerenüberschuss in den chiralen Alkohol der Verbindung (1) umzuwandeln,



wobei das Polypeptid eine Aminosäuresequenz umfasst, die mindestens 80% Sequenzidentität mit SEQ ID NO: 2 und mindestens die folgenden Merkmale aufweist:

- der Rest in der Position, die X206 entspricht, ist I;
- der Rest in der Position, die X40 entspricht, ist R;
- der Rest in der Position, die X153 entspricht, ist I oder L;
- der Rest in der Position, die X190 entspricht, ist A oder P;
- der Rest in der Position, die X196 entspricht, ist T; und
- der Rest in der Position, die X199 entspricht, ist F oder W.

2. Polypeptid nach Anspruch 1, in dem:

(i) die Aminosäuresequenz weiterhin eine Gruppe von Merkmalen umfasst, die aus dem Folgenden ausgewählt ist:

- der Rest in der Position X93 ist A und in der Position X94 ist T;
- der Rest in der Position X93 ist A und in der Position X94 ist S;
- der Rest in der Position X93 ist A und in der Position X94 ist P; oder
- der Rest in der Position X93 ist I und in der Position X94 ist S;

(ii) die Aminosäuresequenz mindestens die folgenden Merkmale umfasst:

der Rest in der Position, die X40 entspricht, ist R;
 der Rest in der Position, die X93 entspricht, ist A;
 der Rest in der Position, die X94 entspricht, ist T;
 der Rest in der Position, die X153 entspricht, ist I;
 der Rest in der Position, die X190 entspricht, ist A;
 der Rest in der Position, die X196 entspricht, ist T;
 der Rest in der Position, die X199 entspricht, ist F;
 der Rest in der Position, die X203 entspricht, ist G; und
 der Rest in der Position, die X206 entspricht, ist I;

(iii) die Aminosäuresequenz die folgenden Merkmale umfasst:

der Rest in der Position, die X40 entspricht, ist R;
 der Rest in der Position, die X93 entspricht, ist A;
 der Rest in der Position, die X94 entspricht, ist S;
 der Rest in der Position, die X96 entspricht, ist N;
 der Rest in der Position, die X153 entspricht, ist L;
 der Rest in der Position, die X190 entspricht, ist A;
 der Rest in der Position, die X196 entspricht, ist T;
 der Rest in der Position, die X199 entspricht, ist F;
 der Rest in der Position, die X202 entspricht, ist G;
 der Rest in der Position, die X203 entspricht, ist G; und
 der Rest in der Position, die X206 entspricht, ist I;

(iv) die Aminosäuresequenz die folgenden Merkmale umfasst:

der Rest in der Position, die X40 entspricht, ist R;
 der Rest in der Position, die X93 entspricht, ist A;
 der Rest in der Position, die X94 entspricht, ist S;
 der Rest in der Position, die X96 entspricht, ist A, P oder N;
 der Rest in der Position, die X153 entspricht, ist L;
 der Rest in der Position, die X190 entspricht, ist A;
 der Rest in der Position, die X196 entspricht, ist T;
 der Rest in der Position, die X199 entspricht, ist F;
 der Rest in der Position, die X202 entspricht, ist G;
 der Rest in der Position, die X203 entspricht, ist G; und
 der Rest in der Position, die X206 entspricht, ist I,

wobei die Aminosäuresequenz gegebenenfalls weiterhin die folgenden Merkmale umfasst:

der Rest in der Position, die X96 entspricht, ist N; und
 der Rest in der Position, die X201 entspricht, ist I oder L;

(v) die Aminosäuresequenz mindestens die folgenden Merkmale umfasst:

der Rest in der Position, die X40 entspricht, ist R;
 der Rest in der Position, die X93 entspricht, ist A;
 der Rest in der Position, die X96 entspricht, ist P;
 der Rest in der Position, die X153 entspricht, ist L;
 der Rest in der Position, die X190 entspricht, ist A;
 der Rest in der Position, die X196 entspricht, ist T;
 der Rest in der Position, die X199 entspricht, ist F;
 der Rest in der Position, die X202 entspricht, ist G;
 der Rest in der Position, die X203 entspricht, ist G; und
 der Rest in der Position, die X206 entspricht, ist I;

wobei die Aminosäuresequenz gegebenenfalls weiterhin ein oder mehrere der folgenden Merkmale umfasst:

der Rest in der Position, die X21 entspricht, ist F;
 der Rest in der Position, die X25 entspricht, ist R;
 der Rest in der Position, die X94 entspricht, ist S;
 der Rest in der Position, die X96 entspricht, ist P;
 5 der Rest in der Position, die X108 entspricht, ist K;
 der Rest in der Position, die X127 entspricht, ist K;
 der Rest in der Position, die X147 entspricht, ist M;
 der Rest in der Position, die X153 entspricht, ist L;
 der Rest in der Position, die X205 entspricht, ist V;
 10 der Rest in der Position, die X211 entspricht, ist K;
 der Rest in der Position, die X221 entspricht, ist D;

(vi) die Aminosäuresequenz mindestens die folgenden Merkmale umfasst:

15 der Rest in der Position, die X21 entspricht, ist F;
 der Rest in der Position, die X25 entspricht, ist R;
 der Rest in der Position, die X40 entspricht, ist R;
 der Rest in der Position, die X93 entspricht, ist A;
 der Rest in der Position, die X94 entspricht, ist S;
 20 der Rest in der Position, die X96 entspricht, ist P;
 der Rest in der Position, die X108 entspricht, ist K;
 der Rest in der Position, die X127 entspricht, ist K;
 der Rest in der Position, die X147 entspricht, ist M;
 der Rest in der Position, die X153 entspricht, ist L;
 25 der Rest in der Position, die X190 entspricht, ist A;
 der Rest in der Position, die X196 entspricht, ist T;
 der Rest in der Position, die X199 entspricht, ist F;
 der Rest in der Position, die X202 entspricht, ist G;
 der Rest in der Position, die X203 entspricht, ist G;
 30 der Rest in der Position, die X206 entspricht, ist I; und
 der Rest in der Position, die X221 entspricht, ist D;

(vii) die Aminosäuresequenz mindestens die folgenden Merkmale umfasst:

35 der Rest in der Position, die X40 entspricht, ist R;
 der Rest in der Position, die X93 entspricht, ist I;
 der Rest in der Position, die X94 entspricht, ist S;
 der Rest in der Position, die X96 entspricht, ist P;
 der Rest in der Position, die X153 entspricht, ist L;
 40 der Rest in der Position, die X190 entspricht, ist A;
 der Rest in der Position, die X196 entspricht, ist T;
 der Rest in der Position, die X199 entspricht, ist F oder W;
 der Rest in der Position, die X201 entspricht, ist A;
 der Rest in der Position, die X202 entspricht, ist G;
 45 der Rest in der Position, die X203 entspricht, ist G;
 der Rest in der Position, die X206 entspricht, ist I; und
 der Rest in der Position, die X211 entspricht, ist K;

wobei die Aminosäuresequenz gegebenenfalls weiterhin ein oder mehrere der folgenden Merkmale umfasst:

50 der Rest in der Position, die X64 entspricht, ist V;
 der Rest in der Position, die X99 entspricht, ist L;
 der Rest in der Position, die X108 entspricht, ist H;
 der Rest in der Position, die X117 entspricht, ist A oder G;
 55 der Rest in der Position, die X147 entspricht, ist I oder M;
 der Rest in der Position, die X195 entspricht, ist M;
 der Rest in der Position, die X205 entspricht, ist V;
 der Rest in der Position, die X223 entspricht, ist I;

der Rest in der Position, die X226 entspricht, ist V;

(viii) die Aminosäuresequenz mindestens die folgenden Merkmale umfasst:

5 der Rest in der Position, die X40 entspricht, ist R;
 der Rest in der Position, die X93 entspricht, ist I;
 der Rest in der Position, die X94 entspricht, ist S;
 der Rest in der Position, die X96 entspricht, ist P;
 der Rest in der Position, die X99 entspricht, ist L;
 10 der Rest in der Position, die X108 entspricht, ist H;
 der Rest in der Position, die X117 entspricht, ist G;
 der Rest in der Position, die X147 entspricht, ist I;
 der Rest in der Position, die X153 entspricht, ist L;
 der Rest in der Position, die X190 entspricht, ist A;
 15 der Rest in der Position, die X196 entspricht, ist T;
 der Rest in der Position, die X199 entspricht, ist F;
 der Rest in der Position, die X201 entspricht, ist A;
 der Rest in der Position, die X202 entspricht, ist G;
 der Rest in der Position, die X203 entspricht, ist G;
 20 der Rest in der Position, die X206 entspricht, ist I;
 der Rest in der Position, die X211 entspricht, ist K; und
 der Rest in der Position, die X226 entspricht, ist V; oder

(ix) die Aminosäuresequenz die folgenden Merkmale umfasst:

25 der Rest in der Position, die X40 entspricht, ist R;
 der Rest in der Position, die X64 entspricht, ist V;
 der Rest in der Position, die X93 entspricht, ist I;
 der Rest in der Position, die X94 entspricht, ist S;
 30 der Rest in der Position, die X96 entspricht, ist P;
 der Rest in der Position, die X99 entspricht, ist L;
 der Rest in der Position, die X108 entspricht, ist H;
 der Rest in der Position, die X117 entspricht, ist G;
 der Rest in der Position, die X147 entspricht, ist I;
 35 der Rest in der Position, die X153 entspricht, ist L;
 der Rest in der Position, die X190 entspricht, ist A;
 der Rest in der Position, die X196 entspricht, ist T;
 der Rest in der Position, die X199 entspricht, ist W;
 der Rest in der Position, die X201 entspricht, ist A;
 40 der Rest in der Position, die X202 entspricht, ist G;
 der Rest in der Position, die X203 entspricht, ist G;
 der Rest in der Position, die X206 entspricht, ist I;
 der Rest in der Position, die X211 entspricht, ist K; und
 der Rest in der Position, die X226 entspricht, ist V.

- 45 3. Polypeptid nach Anspruch 1 oder 2(i), wobei die Aminosäuresequenz weiterhin ein Merkmal oder eine Gruppe von Merkmalen, ausgewählt aus den folgenden, umfasst:

50 der Rest in Position X203 ist G;
 der Rest in Position X202 ist G und in Position X203 ist G; oder
 der Rest in Position X201 ist A, in Position X202 ist G und in Position X203 ist G.

- 55 4. Polypeptid nach einem der Ansprüche 1 bis 3, wobei die Aminosäuresequenz weiterhin ein oder mehrere der folgenden Merkmale umfasst:

 der Rest in der Position, die X21 entspricht, ist R oder F;
 der Rest in der Position, die X25 entspricht, ist R, T oder N;
 der Rest in der Position, die X40 entspricht, ist R;

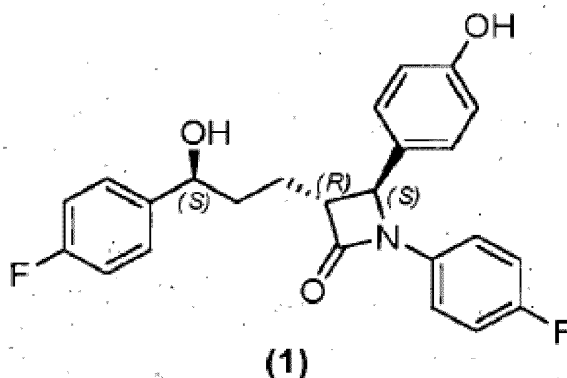
der Rest in der Position, die X64 entspricht, ist V;
 der Rest in der Position, die X93 entspricht, ist A;
 der Rest in der Position, die X94 entspricht, ist T, S oder P;
 der Rest in der Position, die X95 entspricht, ist V oder M;
 5 der Rest in der Position, die X96 entspricht, ist V, G, A, N, S, P oder T;
 der Rest in der Position, die X99 entspricht, ist L;
 der Rest in der Position, die X108 entspricht, ist H;
 der Rest in der Position, die X117 entspricht, ist A oder G;
 der Rest in der Position, die X127 entspricht, ist K oder Q;
 10 der Rest in der Position, die X147 entspricht, ist M oder I;
 der Rest in der Position, die X148 entspricht, ist I;
 der Rest in der Position, die X150 entspricht, ist H oder A;
 der Rest in der Position, die X152 entspricht, ist N oder F;
 der Rest in der Position, die X153 entspricht, ist I oder L;
 15 der Rest in der Position, die X155 entspricht, ist C;
 der Rest in der Position, die X190 entspricht, ist A;
 der Rest in der Position, die X195 entspricht, ist M;
 der Rest in der Position, die X196 entspricht, ist T, A, S, C oder N;
 der Rest in der Position, die X199 entspricht, ist F oder W;
 20 der Rest in der Position, die X201 entspricht, ist I, L oder A;
 der Rest in der Position, die X202 entspricht, ist L, N, V oder G;
 der Rest in der Position, die X203 entspricht, ist G;
 der Rest in der Position, die X204 entspricht, ist V oder A;
 der Rest in der Position, die X205 entspricht, ist V;
 25 der Rest in der Position, die X206 entspricht, ist I;
 der Rest in der Position, die X207 entspricht, ist T, C, I oder N;
 der Rest in der Position, die X211 entspricht, ist K;
 der Rest in der Position, die X221 entspricht, ist D;
 der Rest in der Position, die X223 entspricht, ist I;
 30 der Rest in der Position, die X226 entspricht, ist V.

5. Polypeptid nach Anspruch 1, wobei:

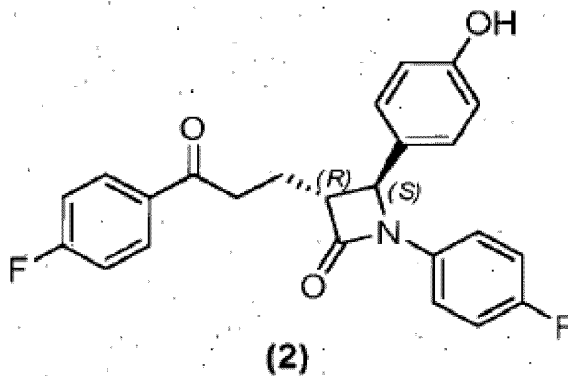
- (i) das Polypeptid fähig ist, die Verbindung (2) mit mindestens 99% Diastereomerenüberschuss in die Verbindung
 35 (1) umzuwandeln;
 (ii) das Polypeptid fähig ist, das Polypeptid zu einer Umwandlung von Verbindung (2) in Verbindung (1) innerhalb
 von 24 h mit einer Substratbeladung von ungefähr 100 g/l von mindestens ungefähr 91%, 92%, 93%, 94%,
 95%, 96%, 97%, 98%, 99% oder mehr fähig ist;
 (iii) das Polypeptid zu einer Umwandlung von Verbindung (2) in Verbindung (1) mit einer im Vergleich zu der
 40 Aktivität des Polypeptids nach SEQ ID NO: 2 um das mindestens 2-fache, mindestens 10-fache, mindestens
 25-fache, mindestens 40-fache oder mindestens 60-fache erhöhten Aktivität fähig ist;
 (iv) das Polypeptid zur Umwandlung von Verbindung (2) in Verbindung (1) innerhalb von 24 h mit einer Subst-
 ratbeladung von ungefähr 100 g/l mit einem Diastereomerenüberschuss von Verbindung (1) von mindestens
 97%, 98%, 99%, 99,1%, 99,2%, 99,3%, 99,4%, 99,5%, 99,6%, 99,7%, 99,8% oder 99,9% oder mehr fähig ist,
 45 wobei die Reaktionsbedingungen gegebenenfalls weiterhin eine IPA-Konzentration von ungefähr 55-75% (v/v),
 einer NADPH- oder NADH-Kofaktorbeladung von ungefähr 0,03-0,5 g/l umfassen und wobei außer dem nicht-
 natürlich vorkommenden Polypeptid kein anderes kofaktorerneuerndes Enzym vorliegt; oder
 (v) das Polypeptid eine Aminosäuresequenz, ausgewählt aus der Gruppe bestehend aus SEQ ID NO: 168, 4,
 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64,
 50 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116,
 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160,
 162, 164 und 166, umfasst.

6. Polynukleotid, das für das Polypeptid nach einem der Ansprüche 1 bis 5 codiert und das gegebenenfalls eine
 55 Nukleotidsequenz, ausgewählt aus der Gruppe bestehend aus SEQ ID NO: 167, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21,
 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81,
 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131,
 133, 135, 137, 139, 141, 143, 145, 147, 149, 151, 153, 155, 157, 159, 161, 163 oder 165, umfasst.

7. Expressionsvektor, umfassend das Polynukleotid nach Anspruch 6, gegebenenfalls umfassend eine Kontrollsequenz, wobei zum Beispiel die Kontrollsequenz einen Promoter umfasst.
8. Wirtszelle, umfassend ein Polynukleotid nach Anspruch 6 oder einen Expressionsvektor nach Anspruch 7, die gegebenenfalls *E. coli* ist.
9. Verfahren zur Herstellung eines nichtnatürlich vorkommenden Polypeptids nach einem der Ansprüche 1 bis 5, umfassend das Kultivieren einer Wirtszelle nach Anspruch 8 und das Isolieren des nichtnatürlich vorkommenden Polypeptids aus der Zelle.
10. Verfahren zur Herstellung von Verbindung (1) oder einem Analog von Verbindung (1) mit Diastereomerenüberschuss,



umfassend das Inkontaktbringen von Verbindung (2) oder einem Analog von Verbindung (2)



unter geeigneten Reaktionsbedingungen mit einem Polypeptid nach einem der Ansprüche 1 bis 5 in Gegenwart von Kofaktor NADPH oder NADH.

11. Verfahren nach Anspruch 10, wobei:

(i) die geeigneten Reaktionsbedingungen Folgendes umfassen:

- (1) Substratbeladung von ungefähr 50-200 g/l Verbindung (2);
- (2) Polypeptidkonzentration von ungefähr 1,5-2,5 g/l;
- (3) NADPH-Kofaktorkonzentration von ungefähr 0,1-0,2 g/l;
- (4) eine Hilfslösungsmittellösung von einem wässrigen Puffer und ungefähr 60-70% (v/v) IPA;
- (5) ungefähr pH 6,6-7,0; und
- (6) Temperatur von ungefähr 28-30°C;

(ii) die geeigneten Reaktionsbedingungen eine Beladung mit Verbindung (2) von mindestens ungefähr 20 g/l, ungefähr 40 g/l, ungefähr 50 g/l, ungefähr 75 g/l, ungefähr 100 g/l, ungefähr 200 g/l, ungefähr 250 g/l, ungefähr 300 g/l oder ungefähr 400 g/l umfassen;

(iii) die geeigneten Reaktionsbedingungen eine Konzentration an modifiziertem Polypeptid von ungefähr 0,1-3,0 g/l, ungefähr 0,5-2,75 g/l, ungefähr 1,0-2,5 g/l, ungefähr 1,5-2,5 g/l, ungefähr 3 g/l, ungefähr 2 g/l, ungefähr 1,5 g/l, ungefähr 1,0 g/l, ungefähr 0,75 g/l oder eine noch niedrigere Konzentration umfassen;

(iv) die geeigneten Reaktionsbedingungen einen pH-Wert von ungefähr 6,0 bis ungefähr 7,5, einen pH-Wert von ungefähr 6,25 bis ungefähr 7,25, einen pH-Wert von ungefähr 6,5 bis ungefähr 7,25, einen pH-Wert von ungefähr 6,6 bis ungefähr 7,25, einen pH-Wert von ungefähr 6,6 bis ungefähr 7,0, einen pH-Wert von ungefähr 6,75 bis ungefähr 7,25 oder einen pH-Wert von ungefähr 6,75 umfassen;

(v) die geeigneten Reaktionsbedingungen eine Temperatur von ungefähr 20°C bis ungefähr 40°C, ungefähr 23°C bis ungefähr 37°C, ungefähr 25°C bis ungefähr 35°C, ungefähr 26°C bis ungefähr 32°C oder ungefähr 28°C bis ungefähr 30°C umfassen;

(vi) die geeigneten Reaktionsbedingungen eine wässrige Pufferlösung, ein organisches Lösungsmittel oder ein Hilfslösungsmittelsystem umfassen;

(vii) die geeigneten Reaktionsbedingungen eine Pufferlösung von ungefähr 0,05 M TEA bis ungefähr 0,25 M TEA umfassen;

(viii) die geeigneten Reaktionsbedingungen ein Hilfslösungsmittelsystem, umfassend ungefähr 30% (v/v) bis ungefähr 70% (v/v) einer wässrigen Pufferlösung und ungefähr 30% (v/v) bis ungefähr 70% (v/v) einer organischen Lösungsmittellösung umfassen, wobei das Hilfslösungsmittelsystem gegebenenfalls eine wässrige Pufferlösung, IPA und ein anderes organisches Lösungsmittel umfasst wie zum Beispiel wobei:

(a) das Hilfslösungsmittelsystem ungefähr 45-55% (v/v) einer wässrigen Pufferlösung, ungefähr 15-25% (v/v) IPA und ungefähr 25-35% (v/v) Toluol umfasst; oder

(b) das Hilfslösungsmittelsystem ungefähr 50% (v/v) einer wässrigen Pufferlösung, ungefähr 20% (v/v) IPA und ungefähr 30% (v/v) Toluol umfasst;

(ix) die geeigneten Reaktionsbedingungen ein Hilfslösungsmittelsystem, umfassend eine wässrige Pufferlösung und IPA, umfassen, wobei die IPA-Konzentration ungefähr 25-75% (v/v), ungefähr 35-75% (v/v), ungefähr 45-75% (v/v), ungefähr 55-75% (v/v), ungefähr 60-70% (v/v), ungefähr 62-68% (v/v) beträgt, wobei gegebenenfalls das Hilfslösungsmittelsystem 0,1 M TEA-Puffer und ungefähr 60% (v/v) bis ungefähr 70% (v/v) IPA umfasst, wobei das Hilfslösungsmittelsystem zum Beispiel ungefähr 35% (v/v) 0,1 M TEA-Puffer und ungefähr 65% (v/v) IPA umfasst;

(x) die geeigneten Reaktionsbedingungen eine NADH- oder NADPH-Kofaktorkonzentration von ungefähr 0,03-0,5 g/l, ungefähr 0,05-0,3 g/l, ungefähr 0,1-0,2 g/l, ungefähr 0,5 g/l, ungefähr 0,1 g/l oder ungefähr 0,2 g/l umfassen;

(xi) das Verfahren weiterhin umfasst, dass man NADP⁺ mit einem Kofaktorerneuerungssystem in NADPH umwandelt, wobei gegebenenfalls das Kofaktorerneuerungssystem aus einer Ketoreduktase und einem sekundären Alkohol, oder Glukosedehydrogenase und Glukose, ausgewählt ist, wobei zum Beispiel:

(a) der sekundäre Alkohol IPA ist; oder

(b) die Ketoreduktase das Polypeptid nach einem der Ansprüche 1 bis 5 ist;

(xii) die geeigneten Reaktionsbedingungen eine IPA-Konzentration von ungefähr 55-75% (v/v), eine NADPH- oder NADH-Kofaktorbiladung von ungefähr 0,03-0,5 g/l umfassen und wobei außer dem Polypeptid nach einem der Ansprüche 1 bis 5 kein anderes Kofaktorerneuerungsenzym vorliegt;

(xiii) die geeigneten Reaktionsbedingungen Folgendes umfassen:

(1) eine Substratbeladung von ungefähr 100 g/l Verbindung (2);

(2) eine Polypeptidkonzentration von ungefähr 2,0 g/l;

(3) eine NADPH-Kofaktorkonzentration von ungefähr 0,1 g/l;

(4) eine Hilfslösungsmittellösung von einem wässrigen Puffer von 0,1 M TEA und ungefähr 65% (v/v) IPA;

(5) einen pH-Wert von ungefähr 6,75; und

(6) eine Temperatur von ungefähr 30°C;

(xiv) die geeigneten Reaktionsbedingungen umfassen:

(1) eine Substratbeladung von ungefähr 50-150 g/l Verbindung (2);

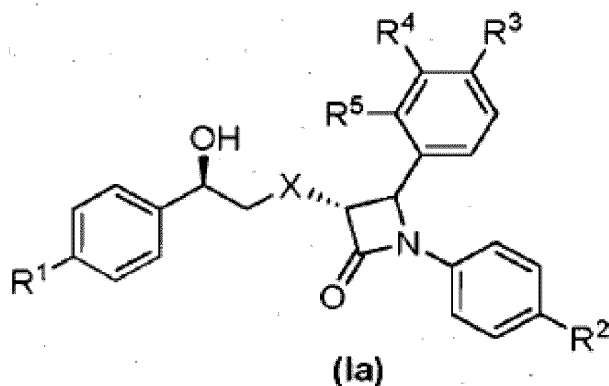
- (2) eine Polypeptidkonzentration von ungefähr 2,5-3,5 g/l;
 (3) eine NADPH-Kofaktorkonzentration von ungefähr 0,1-0,2 g/l;
 (4) eine Hilfslösungsmittellösung von einem wässrigen Puffer und ungefähr 15-25% (v/v) IPA und 25-35% (v/v) Toluol;
 (5) einen pH-Wert von ungefähr 6,6-7,0; und
 (6) eine Temperatur von ungefähr 28-30°C;

(xv) die geeigneten Reaktionsbedingungen umfassen:

- (1) eine Substratbeladung von ungefähr 100 g/l Verbindung (2);
 (2) eine Polypeptidkonzentration von ungefähr 3,0 g/l;
 (3) eine NADPH-Kofaktorkonzentration von ungefähr 0,1 g/l;
 (4) eine Hilfslösungsmittellösung von einem wässrigen Puffer von 0,1 M TEA und ungefähr 20% (v/v) IPA und 30% (v/v) Toluol;
 (5) einen pH-Wert von ungefähr 6,6-7,0; und
 (6) eine Temperatur von ungefähr 28-30°C;

(xvi) die geeigneten Reaktionsbedingungen eine Substratbeladung von ungefähr 100 g/l Verbindung (2) umfassen und wobei das Verfahren zu einer Umwandlung von Verbindung (2) in Verbindung (1) innerhalb von 24 h von mindestens ungefähr 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% oder mehr führt; oder
 (xvii) die geeigneten Reaktionsbedingungen eine Substratbeladung von ungefähr 100 g/l Verbindung (2) umfassen und wobei das Verfahren innerhalb von 24 h zu einem Diastereomerenüberschuss von Verbindung (1) von mindestens 97%, 98%, 99%, 99,1%, 99,2%, 99,3%, 99,4%, 99, 5%, 99,6%, 99,7%, 99,8% oder 99,9% oder mehr führt.

12. Verfahren nach Anspruch 10, wobei es sich bei dem Analog von Verbindung (1) um eine Verbindung der Formel Ia handelt:



worin

X C oder S bedeutet;

R¹ aus -H, -F, -Cl, -Br oder -I ausgewählt ist;

R² aus -H, -F, -Cl, -Br, -I, -CN, -OH (das gegebenenfalls mit einer Hydroxylschutzgruppe geschützt ist), -CH₂NH₂ (das gegebenenfalls mit einer Stickstoffschutzgruppe geschützt ist) und einer der folgenden gegebenenfalls substituierten Gruppen: Alkyl, Alkoxy, Alkenyl, Alkenoxy, Alkinyl, Alkinoxy, Cycloalkyl, Aryl, Heteroaryl oder Heterocyclus ausgewählt ist;

R³ aus -H, -F, -Cl, -Br, -I, -CN, -OH (das gegebenenfalls mit einer Hydroxylschutzgruppe geschützt ist), -CH₂NH₂ (das gegebenenfalls mit einer Stickstoffschutzgruppe geschützt ist) und einer der folgenden gegebenenfalls substituierten Gruppen: Alkyl, Alkoxy, Alkenyl, Alkenoxy, Alkinyl, Alkinoxy, Cycloalkyl, Aryl, Heteroaryl oder Heterocyclus ausgewählt ist;

R⁴ aus -H, -F, -Cl, -Br, -I, -CN, -OH (das gegebenenfalls mit einer Hydroxylschutzgruppe geschützt ist) ausgewählt ist; und

R⁵ aus -H, -F, -Cl, -Br, -I, -CN, -OH (das gegebenenfalls mit einer Hydroxylschutzgruppe geschützt ist), -CH₂NH₂ (das gegebenenfalls mit einer Stickstoffschutzgruppe geschützt ist) und einer der folgenden gegebenenfalls substituierten Gruppen: Alkyl, Alkoxy, Alkenyl, Alkenoxy, Alkynyl, Alkinoxy, Cycloalkyl, Aryl, Heteroaryl oder Heterocyclus ausgewählt ist.

13. Verfahren nach Anspruch 12, wobei:

(i) R² und R³ jeweils unabhängig aus -H, -F, -Cl, -Br, -I, -CN, -OH (das gegebenenfalls mit einer Hydroxylschutzgruppe geschützt ist), -CH₂NH₂ (das gegebenenfalls mit einer Stickstoffschutzgruppe geschützt ist) und einer beliebigen der folgenden gegebenenfalls substituierten Gruppen: -(C₁-C₆)-Alkyl, -(C₁-C₆)-Alkoxy, -(C₁-C₆)-Alkenyl, -(C₁-C₆)-Alkenoxy, -(C₁-C₆)-Alkynyl, -(C₁-C₆)-Alkinoxy, -(C₁-C₆)-Cycloalkyl oder einem Heterocyclus mit 1 bis 4 Kohlenstoffatomen und 1 bis 2 Heteroatomen ausgewählt sind,

(ii) R² und R³ jeweils unabhängig aus -H, -F, -Cl, -Br, -I, -CN, -OH (das gegebenenfalls mit einer Hydroxylschutzgruppe geschützt ist), -CH₂NH₂ (das gegebenenfalls mit einer Stickstoffschutzgruppe geschützt ist) und einer beliebigen der folgenden gegebenenfalls substituierten Gruppen: -(C₁-C₄)-Alkyl, -(C₁-C₄)-Alkoxy, -(C₁-C₄)-Alkenyl, -(C₁-C₄)-Alkenoxy, -(C₁-C₄)-Alkynyl, -(C₁-C₄)-Alkinoxy, -(C₁-C₄)-Cycloalkyl oder einem Heterocyclus mit 1 bis 3 Kohlenstoffatomen und 1 bis 2 Heteroatomen ausgewählt sind,

(iii) R² und R³ jeweils unabhängig aus -H, -F, -Cl, -Br, -I, -CN, -OH (das gegebenenfalls mit einer Hydroxylschutzgruppe geschützt ist) und -CH₂NH₂ (das gegebenenfalls mit einer Stickstoffschutzgruppe geschützt ist) ausgewählt sind;

(iv) R² und R³ jeweils unabhängig aus den folgenden gegebenenfalls substituierten Gruppen ausgewählt sind: -(C₁-C₄)-Alkyl, -(C₁-C₄)-Alkoxy, -(C₁-C₄)-Alkenyl, -(C₁-C₄)-Alkenoxy, -(C₁-C₄)-Alkynyl und -(C₁-C₄)-Alkinoxy;

(v) R² und R³ gegebenenfalls mit einer oder mehreren -OH-Gruppen substituiert sind; oder

(vi) ein beliebiges oder mehrere von R², R³ oder R⁵ eine -OH-Gruppe oder eine -OH-Gruppe, die mit einer Hydroxylschutzgruppe geschützt ist, bedeutet, wobei die Hydroxylschutzgruppe gegebenenfalls aus der Gruppe bestehend aus Benzyl, Acetyl, Benzoyl, tert.-Butyloxycarbonyl, Silyl, tert.-Butyldiphenylsilyl, Trimethylsilyl, para-Methoxybenzyl, Benzyliden, Dimethylacetal und Methoxymethyl ausgewählt ist, wobei zum Beispiel die Silylgruppe -Si-(R^a)(R^b)(R^c) ist und R^a, R^b und R^c jeweils unabhängig aus der Gruppe bestehend aus C₁-C₆-Alkyl-, Phenyl-, Acetyl- und Benzylgruppen ausgewählt sind.

14. Verfahren nach Anspruch 12 oder Anspruch 13, wobei

(i) R¹ -F bedeutet, R⁴ -H bedeutet und R⁵ -H bedeutet;

(ii) R² -F bedeutet; und/oder

(iii) X C oder S bedeutet.

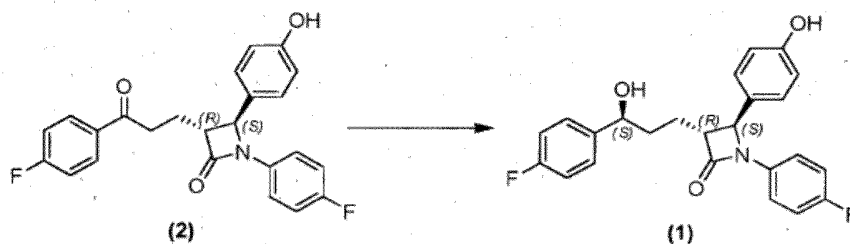
15. Verfahren nach einem der Ansprüche 10 bis 14, wobei:

(i) das Verfahren weiterhin einen oder mehrere Schritte, ausgewählt aus den folgenden, umfasst: Extrahieren von Verbindung (1); Isolieren von Verbindung (1); Aufreinigen von Verbindung (1); und Kristallisieren von Verbindung (1); und/oder

(ii) das Polypeptid eine Aminosäuresequenz, ausgewählt aus der Gruppe bestehend aus SEQ ID NO: 168, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164 und 166, umfasst.

Revendications

1. Polypeptide d'origine non naturelle capable de convertir la cétone de composé (2) en alcool chiral de composé (1) en excès diastéréoisomérique



le polypeptide comprenant une séquence d'acides aminés qui a au moins 80 % d'identité de séquence à SEQ ID NO: 2 et au moins les caractéristiques suivantes :

- le résidu à la position correspondant à X206 est I ;
- le résidu à la position correspondant à X40 est R ;
- le résidu à la position correspondant à X153 est I ou L ;
- le résidu à la position correspondant à X190 est A ou P ;
- le résidu à la position correspondant à X196 est T ; et
- le résidu à la position correspondant à X199 est F ou W.

2. Polypeptide de la revendication 1 dans lequel :

- (i) la séquence d'acides aminés comprend en outre un groupe de caractéristiques choisies parmi :

- le résidu à la position X93 est A et à la position X94 est T ;
- le résidu à la position X93 est A et à la position X94 est S ;
- le résidu à la position X93 est A et à la position X94 est P ; ou
- le résidu à la position X93 est I et à la position X94 est S ;

- (ii) la séquence d'acides aminés comprend au moins les caractéristiques suivantes :

- le résidu à la position correspondant à X40 est R ;
- le résidu à la position correspondant à X93 est A ;
- le résidu à la position correspondant à X94 est T ;
- le résidu à la position correspondant à X153 est I ;
- le résidu à la position correspondant à X190 est A ;
- le résidu à la position correspondant à X196 est T ;
- le résidu à la position correspondant à X199 est F ;
- le résidu à la position correspondant à X203 est G ; et
- le résidu à la position correspondant à X206 est I ;

- (iii) la séquence d'acides aminés comprend les caractéristiques suivantes :

- le résidu à la position correspondant à X40 est R ;
- le résidu à la position correspondant à X93 est A ;
- le résidu à la position correspondant à X94 est S ;
- le résidu à la position correspondant à X96 est N ;
- le résidu à la position correspondant à X153 est L ;
- le résidu à la position correspondant à X190 est A ;
- le résidu à la position correspondant à X196 est T ;
- le résidu à la position correspondant à X199 est F ;
- le résidu à la position correspondant à X202 est G ;
- le résidu à la position correspondant à X203 est G ; et
- le résidu à la position correspondant à X206 est I ;

- (iv) la séquence d'acides aminés comprend les caractéristiques suivantes :

EP 2 566 497 B1

le résidu à la position correspondant à X40 est R ;
le résidu à la position correspondant à X93 est A ;
le résidu à la position correspondant à X94 est S ;
le résidu à la position correspondant à X96 est A, P ou
N ;
le résidu à la position correspondant à X153 est L ;
le résidu à la position correspondant à X190 est A ;
le résidu à la position correspondant à X196 est T ;
le résidu à la position correspondant à X199 est F ;
le résidu à la position correspondant à X202 est G ;
le résidu à la position correspondant à X203 est G ; et
le résidu à la position correspondant à X206 est I,

facultativement dans lequel la séquence d'acides aminés comprend en outre les caractéristiques suivantes :

le résidu à la position correspondant à X96 est N ; et
le résidu à la position correspondant à X201 est I ou L ;

(v) la séquence d'acides aminés comprend au moins les caractéristiques suivantes :

le résidu à la position correspondant à X40 est R ;
le résidu à la position correspondant à X93 est A ;
le résidu à la position correspondant à X96 est P ;
le résidu à la position correspondant à X153 est L ;
le résidu à la position correspondant à X190 est A ;
le résidu à la position correspondant à X196 est T ;
le résidu à la position correspondant à X199 est F ;
le résidu à la position correspondant à X202 est G ;
le résidu à la position correspondant à X203 est G ; et
le résidu à la position correspondant à X206 est I,

facultativement dans lequel la séquence d'acides aminés comprend en outre une ou plusieurs des caractéristiques suivantes :

le résidu à la position correspondant à X21 est F ;
le résidu à la position correspondant à X25 est R ;
le résidu à la position correspondant à X94 est S ;
le résidu à la position correspondant à X96 est P ;
le résidu à la position correspondant à X108 est K ;
le résidu à la position correspondant à X127 est K ;
le résidu à la position correspondant à X147 est M ;
le résidu à la position correspondant à X153 est L ;
le résidu à la position correspondant à X205 est V ;
le résidu à la position correspondant à X211 est K ;
le résidu à la position correspondant à X221 est D ;

(vi) la séquence d'acides aminés comprend au moins les caractéristiques suivantes :

le résidu à la position correspondant à X21 est F ;
le résidu à la position correspondant à X25 est R ;
le résidu à la position correspondant à X40 est R ;
le résidu à la position correspondant à X93 est A ;
le résidu à la position correspondant à X94 est S ;
le résidu à la position correspondant à X96 est P ;
le résidu à la position correspondant à X108 est K ;
le résidu à la position correspondant à X127 est K ;
le résidu à la position correspondant à X147 est M ;
le résidu à la position correspondant à X153 est L ;

5 le résidu à la position correspondant à X190 est A ;
le résidu à la position correspondant à X196 est T ;
le résidu à la position correspondant à X199 est F ;
le résidu à la position correspondant à X202 est G ;
le résidu à la position correspondant à X203 est G ;
le résidu à la position correspondant à X206 est I et
le résidu à la position correspondant à X221 est D ;

10 (vii) la séquence d'acides aminés comprend au moins les caractéristiques suivantes :

le résidu à la position correspondant à X40 est R ;
le résidu à la position correspondant à X93 est I ;
le résidu à la position correspondant à X94 est S ;
le résidu à la position correspondant à X96 est P ;
15 le résidu à la position correspondant à X153 est L ;
le résidu à la position correspondant à X190 est A ;
le résidu à la position correspondant à X196 est T ;
le résidu à la position correspondant à X199 est F ou W ;
le résidu à la position correspondant à X201 est A ;
20 le résidu à la position correspondant à X202 est G ;
le résidu à la position correspondant à X203 est G ;
le résidu à la position correspondant à X206 est I ; et
le résidu à la position correspondant à X211 est K,

25 facultativement dans lequel la séquence d'acides aminés comprend en outre une ou plusieurs des caractéristiques suivantes :

le résidu à la position correspondant à X64 est V ;
le résidu à la position correspondant à X99 est L ;
30 le résidu à la position correspondant à X108 est H ;
le résidu à la position correspondant à X117 est A ou G ;
le résidu à la position correspondant à X147 est I ou M ;
le résidu à la position correspondant à X195 est M ;
le résidu à la position correspondant à X205 est V ;
35 le résidu à la position correspondant à X223 est I ;
le résidu à la position correspondant à X226 est V ;

(viii) la séquence d'acides aminés comprend au moins les caractéristiques suivantes :

40 le résidu à la position correspondant à X40 est R ;
le résidu à la position correspondant à X93 est I ;
le résidu à la position correspondant à X94 est S ;
le résidu à la position correspondant à X96 est P ;
le résidu à la position correspondant à X99 est L ;
45 le résidu à la position correspondant à X108 est H ;
le résidu à la position correspondant à X117 est G ;
le résidu à la position correspondant à X147 est I ;
le résidu à la position correspondant à X153 est L ;
le résidu à la position correspondant à X190 est A ;
50 le résidu à la position correspondant à X196 est T ;
le résidu à la position correspondant à X199 est F ;
le résidu à la position correspondant à X201 est A ;
le résidu à la position correspondant à X202 est G ;
le résidu à la position correspondant à X203 est G ;
55 le résidu à la position correspondant à X206 est I ;
le résidu à la position correspondant à X211 est K ; et
le résidu à la position correspondant à X226 est V ; ou

(ix) la séquence d'acides aminés comprend les caractéristiques suivantes :

5 le résidu à la position correspondant à X40 est R ;
le résidu à la position correspondant à X64 est V ;
le résidu à la position correspondant à X93 est I ;
le résidu à la position correspondant à X94 est S ;
le résidu à la position correspondant à X96 est P ;
le résidu à la position correspondant à X99 est L ;
10 le résidu à la position correspondant à X108 est H ;
le résidu à la position correspondant à X117 est G ;
le résidu à la position correspondant à X147 est I ;
le résidu à la position correspondant à X153 est L ;
le résidu à la position correspondant à X190 est A ;
le résidu à la position correspondant à X196 est T ;
15 le résidu à la position correspondant à X199 est W ;
le résidu à la position correspondant à X201 est A ;
le résidu à la position correspondant à X202 est G ;
le résidu à la position correspondant à X203 est G ;
le résidu à la position correspondant à X206 est I ;
20 le résidu à la position correspondant à X211 est K ; et
le résidu à la position correspondant à X226 est V.

3. Polypeptide de la revendication 1 ou 2(i) dans lequel la séquence d'acides aminés comprend en outre une caractéristique ou un groupe de caractéristiques choisies parmi :

25 le résidu à la position X203 est G ;
le résidu à la position X202 est G et à la position X203 est G ; ou
le résidu à la position X201 est A, à la position X202 est G et à la position X203 est G.

30 4. Polypeptide de l'une quelconque des revendications 1 à 3 dans lequel la séquence d'acides aminés comprend en outre une ou plusieurs des caractéristiques suivantes :

le résidu à la position correspondant à X21 est R ou F ;
le résidu à la position correspondant à X25 est R, T ou N ;
35 le résidu à la position correspondant à X40 est R ;
le résidu à la position correspondant à X64 est V ;
le résidu à la position correspondant à X93 est A ;
le résidu à la position correspondant à X94 est T, S ou P ;
le résidu à la position correspondant à X95 est V ou M ;
40 le résidu à la position correspondant à X96 est V, G, A, N, S, P ou T ;
le résidu à la position correspondant à X99 est L ;
le résidu à la position correspondant à X108 est H ;
le résidu à la position correspondant à X117 est A ou G ;
le résidu à la position correspondant à X127 est K ou Q ;
45 le résidu à la position correspondant à X147 est M ou I ;
le résidu à la position correspondant à X148 est I ;
le résidu à la position correspondant à X150 est H ou A ;
le résidu à la position correspondant à X152 est N ou F ;
le résidu à la position correspondant à X153 est I ou L ;
50 le résidu à la position correspondant à X155 est C ;
le résidu à la position correspondant à X190 est A ;
le résidu à la position correspondant à X195 est M ;
le résidu à la position correspondant à X196 est T, A, S, C ; ou N ;
le résidu à la position correspondant à X199 est F ou W ;
55 le résidu à la position correspondant à X201 est I, L ou A ;
le résidu à la position correspondant à X202 est L, N, V ou G ;
le résidu à la position correspondant à X203 est G ;
le résidu à la position correspondant à X204 est V ou A ;

le résidu à la position correspondant à X205 est V ;
 le résidu à la position correspondant à X206 est I ;
 le résidu à la position correspondant à X207 est T, C, I ou N ;
 le résidu à la position correspondant à X211 est K ;
 le résidu à la position correspondant à X221 est D ;
 le résidu à la position correspondant à X223 est I ;
 le résidu à la position correspondant à X226 est V.

5. Polypeptide de la revendication 1 **caractérisé en ce que** :

- (i) le polypeptide est capable de convertir le composé (2) en composé (1) avec un excès diastéréoisomérique d'au moins 99 % ;
- (ii) le polypeptide est capable d'effectuer la conversion d'au moins environ 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 % ou plus du composé (2) en composé (1) en 24 h avec une charge de substrat d'environ 100 g/l ;
- (iii) le polypeptide est capable de convertir le composé (2) en composé (1) avec une activité augmentée au moins 2 fois, au moins 10 fois, au moins 25 fois, au moins 40 fois ou au moins 60 fois par rapport à l'activité du polypeptide de SEQ ID NO: 2 ;
- (iv) le polypeptide est capable de convertir le composé (2) en composé (1) avec un excès diastéréoisomérique de composé (1) d'au moins 97 %, 98, 99 %, 99,1 %, 99,2 %, 99,3 %, 99,4 %, 99,5 %, 99,6 %, 99,7 %, 99,8 % ou 99,9 % ou plus en 24 h avec une charge de substrat d'environ 100 g/l, facultativement **caractérisé en ce que** les conditions de réaction comprennent en outre une concentration d'IPA d'environ 55 à 75 % (v/v), une charge de cofacteur NADPH ou NADH d'environ 0,03 à 0,5 g/l, et où aucune enzyme de recyclage de cofacteur n'est présente autre que le polypeptide d'origine non naturelle ; ou
- (v) le polypeptide comprend une séquence d'acides aminés choisie dans le groupe constitué de SEQ ID NO: 168, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164 et 166.

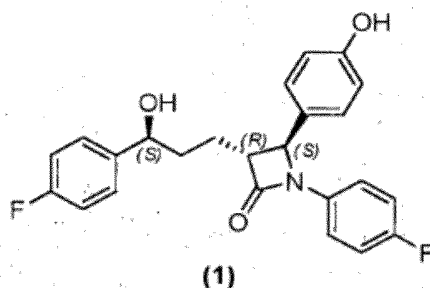
6. Polynucléotide codant pour le polypeptide de l'une quelconque des revendications 1 à 5, qui comprend facultativement une séquence nucléotidique choisie dans le groupe constitué de SEQ ID NO: 167, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139, 141, 143, 145, 147, 149, 151, 153, 155, 157, 159, 161, 163 ou 165.

7. Vecteur d'expression comprenant le polynucléotide de la revendication 6, comprenant facultativement une séquence de contrôle, par exemple, dans lequel la séquence de contrôle comprend un promoteur.

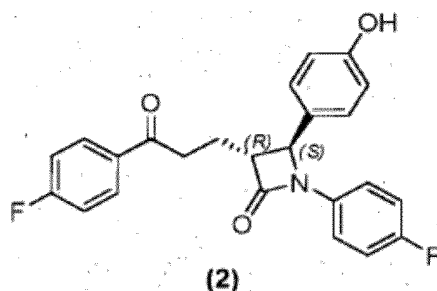
8. Cellule hôte comprenant un polynucléotide de la revendication 6 ou un vecteur d'expression de la revendication 7, qui est facultativement *E. coli*.

9. Procédé de préparation d'un polypeptide d'origine non naturelle de l'une quelconque des revendications 1 à 5 comprenant la culture d'une cellule hôte de la revendication 8 et l'isolement du polypeptide d'origine non naturelle à partir de la cellule.

10. Procédé de préparation du composé (1) ou d'un analogue du composé (1) en excès diastéréoisomérique



comprenant la mise en contact du composé (2) ou d'un analogue du composé (2)



avec un polypeptide de l'une quelconque des revendications 1 à 5 en présence de cofacteur NADPH ou NADH dans des conditions de réaction adaptées.

11. Procédé de la revendication 10 dans lequel :

(i) les conditions de réaction adaptées comprennent :

- (1) une charge de substrat d'environ 50 à 200 g/l de composé (2) ;
- (2) une concentration de polypeptide d'environ 1,5 à 2,5 g/l ;
- (3) une concentration de cofacteur NADPH d'environ 0,1 à 0,2 g/l ;
- (4) une solution de co-solvant d'un tampon aqueux et environ 60 à 70 % (v/v) d'IPA ;
- (5) un pH d'environ 6,6 à 7,0 ; et
- (6) une température d'environ 28 à 30 °C ;

(ii) les conditions de réaction adaptées comprennent une charge de composé (2) d'au moins environ 20 g/l, environ 40 g/l, environ 50 g/l, environ 75 g/l, environ 100 g/l, environ 200 g/l, environ 250 g/l, environ 300 g/l ou environ 400 g/l ;

(iii) les conditions de réaction adaptées comprennent une concentration de polypeptide modifié d'environ 0,1 à 3,0 g/l, environ 0,5 à 2,75 g/l, environ 1,0 à 2,5 g/l, environ 1,5 à 2,5 g/l, environ 3 g/l, environ 2 g/l, environ 1,5 g/l, environ 1,0 g/l, environ 0,75 g/l ou même une concentration plus faible ;

(iv) les conditions de réaction adaptées comprennent un pH d'environ 6,0 à environ 7,5, un pH d'environ 6,25 à environ 7,25, un pH d'environ 6,5 à environ 7,25, un pH d'environ 6,6 à environ 7,25, un pH d'environ 6,6 à environ 7,0, un pH d'environ 6,75 à environ 7,25 ou un pH d'environ 6,75 ;

(v) les conditions de réaction adaptées comprennent une température d'environ 20 °C à environ 40 °C, environ 23 °C à environ 37 °C, environ 25 °C à environ 35 °C, environ 26 °C à environ 32 °C ou environ 28 °C à environ 30 °C ;

(vi) les conditions de réaction adaptées comprennent une solution tampon aqueuse, un solvant organique ou un système de co-solvant ;

(vii) les conditions de réaction adaptées comprennent une solution tampon d'environ 0,05 M TEA à environ 0,25 M TEA ;

(viii) les conditions de réaction adaptées comprennent un système de co-solvant comprenant environ 30 % (v/v) à environ 70 % (v/v) d'une solution tampon aqueuse et environ 30 % (v/v) à environ 70 % (v/v) d'une

solution organique de solvant, facultativement dans lequel le système de co-solvant comprend une solution tampon aqueuse, IPA, et un autre solvant organique, par exemple dans lequel :

- 5 (a) le système de co-solvant comprend environ 45 à 55 % (v/v) d'une solution tampon aqueuse, environ 15 à 25 % (v/v) d'IPA, et environ 25 à 35 % (v/v) de toluène ; ou
(b) le système de co-solvant comprend environ 50 % (v/v) d'une solution tampon aqueuse, environ 20 % (v/v) d'IPA, et environ 30 % (v/v) de toluène ;

- 10 (ix) les conditions de réaction adaptées comprennent un système de co-solvant comprenant une solution tampon aqueuse et IPA, la concentration d'IPA étant d'environ 25 à 75 % (v/v), environ 35 à 76 % (v/v), environ 45 à 75 % (v/v), environ 55 à 75 % (v/v), environ 60 à 70 % (v/v), environ 62 à 68 % (v/v), facultativement dans lequel le système de co-solvant comprend un tampon TEA 0,1 M et environ 60 % (v/v) à environ 70 % (v/v) d'IPA, par exemple dans lequel le système de co-solvant comprend environ 35 % (v/v) de tampon TEA 0,1 M et environ 65 % (v/v) d'IPA ;

- 15 (x) les conditions de réaction adaptées comprennent une concentration de cofacteur NADH ou NADPH d'environ 0,03 à 0,5 g/l, environ 0,05 à 0,3 g/l, environ 0,1 à 0,2 g/l, environ 0,5 g/l, environ 0,1 g/l ou environ 0,2 g/l ;

- (xi) le procédé comprend en outre la conversion de NADP⁺ en NADPH avec un système de recyclage de cofacteur, facultativement dans lequel le système de recyclage de cofacteur est choisi parmi une cétoréductase et un alcool secondaire ou une glucose déshydrogénase et du glucose, par exemple dans lequel :

- 20 (a) l'alcool secondaire est IPA ; ou
(b) la cétoréductase est le polypeptide de l'une quelconque des revendications 1 à 5 ;

- 25 (xii) les conditions de réaction adaptées comprennent une concentration d'IPA d'environ 55 à 75 % (v/v), une charge de cofacteur NADPH ou NADH d'environ 0,03 à 0,5 g/l, et dans lequel aucune enzyme de recyclage de cofacteur n'est présente autre que le polypeptide de l'une quelconque des revendications 1 à 5 ;

- (xiii) les conditions de réaction adaptées comprennent :

- 30 (1) une charge de substrat d'environ 100 g/l de composé (2) ;
(2) une concentration de polypeptide d'environ 2,0 g/l ;
(3) une concentration de cofacteur NADPH d'environ 0,1 g/l ;
(4) une solution de co-solvant d'un tampon aqueux de TEA 0,1 M et environ 65 % (v/v) d'IPA ;
(5) un pH d'environ 6,75 ; et
(6) une température d'environ 30 °C ;

- 35 (xiv) les conditions de réaction adaptées comprennent :

- 40 (1) une charge de substrat d'environ 50 à 150 g/l de composé (2) ;
(2) une concentration de polypeptide d'environ 2,5 à 3,5 g/l ;
(3) une concentration de cofacteur NADPH d'environ 0,1 à 0,2 g/l ;
(4) une solution de co-solvant d'un tampon aqueux de et d'environ 15 à 25 % (v/v) d'IPA et 25 à 35 % (v/v) de toluène ;
(5) un pH d'environ 6,6 à 7,0 ; et
(6) une température d'environ 28 à 30 °C ;

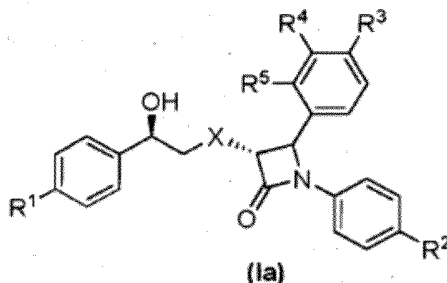
- 45 (xv) les conditions de réaction adaptées comprennent :

- 50 (1) une charge de substrat d'environ 100 g/l de composé (2) ;
(2) une concentration de polypeptide d'environ 3,0 g/l ;
(3) une concentration de cofacteur NADPH d'environ 0,1 g/l ;
(4) une solution de co-solvant d'un tampon aqueux de TEA 0,1 M et environ 20 % (v/v) d'IPA et 30 % (v/v) de toluène ;
(5) un pH d'environ 6,6 à 7,0 ; et
(6) une température d'environ 28 à 30 °C ;

- 55 (xvi) les conditions de réaction adaptées comprennent une charge de substrat d'environ 100 g/l de composé (2) et le procédé conduisant à au moins environ 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 % ou plus de conversion du composé (2) en composé (1) en 24 h ; ou

(xvii) les conditions de réaction adaptées comprennent une charge de substrat d'environ 100 g/l de composé (2) et le procédé conduisant à un excès diastéréoisomérique de composé (1) d'au moins 97 %, 98, 99 %, 99,1 %, 99,2 %, 99,3 %, 99,4 %, 99,5 %, 99,6 %, 99,7 %, 99,8 % ou 99,9 % ou plus en 24 h.

12. Procédé de la revendication 10 dans lequel l'analogue de composé (1) est un composé de formule la :



dans laquelle,

X est C ou S ;

R¹ est choisi parmi -H, -F, -Cl, -Br ou -I ;

R² est choisi parmi -H, -F, -Cl, -Br, -I, -CN, -OH (facultativement protégé avec un groupe protecteur d'hydroxyle), -CH₂NH₂ (facultativement protégé avec un groupe protecteur d'azote), et l'un quelconque des groupes facultativement substitués suivants : alkyle, alcoxy, alcényle, alcénoxy, alcynyle, alcynoxy, cycloalkyle, aryle, hétéroaryle ou hétérocycle ;

R³ est choisi parmi -H, -F, -Cl, -Br, -I, -CN, -OH (facultativement protégé avec un groupe protecteur d'hydroxyle), -CH₂NH₂ (facultativement protégé avec un groupe protecteur d'azote), et l'un quelconque des groupes facultativement substitués suivants : alkyle, alcoxy, alcényle, alcénoxy, alcynyle, alcynoxy, cycloalkyle, aryle, hétéroaryle ou hétérocycle ;

R⁴ est choisi parmi -H, -F, -Cl, -Br, -I, -CN, -OH (facultativement protégé avec un groupe protecteur d'hydroxyle) ;
et

R⁵ est choisi parmi -H, -F, -Cl, -Br, -I, -CN, -OH (facultativement protégé avec un groupe protecteur d'hydroxyle), -CH₂NH₂ (facultativement protégé avec un groupe protecteur d'azote), et l'un quelconque des groupes facultativement substitués suivants : alkyle, alcoxy, alcényle, alcénoxy, alcynyle, alcynoxy, cycloalkyle, aryle, hétéroaryle ou hétérocycle.

13. Procédé de la revendication 12 dans lequel :

(i) R² et R³ sont chacun indépendamment choisis parmi -H, -F, -Cl, -Br, -I, -CN, -OH (facultativement protégé avec un groupe protecteur d'hydroxyle), -CH₂NH₂ (facultativement protégé avec un groupe protecteur d'azote), et l'un quelconque des groupes facultativement substitués suivants : -(alkyle en C₁-C₆), -(alcoxy en C₁-C₆), -(alcényle en C₁-C₆), -(alcénoxy en C₁-C₆), -(alcynyle en C₁-C₆), -(alcynoxy en C₁-C₆), -(cycloalkyle en C₁-C₆) ou un hétérocycle, ayant de 1 à 4 atomes de carbone et 1 à 2 hétéroatomes ;

(ii) R² et R³ sont chacun indépendamment choisis parmi -H, -F, -Cl, -Br, -I, -CN, -OH (facultativement protégé avec un groupe protecteur d'hydroxyle), -CH₂NH₂ (facultativement protégé avec un groupe protecteur d'azote), et l'un quelconque des groupes facultativement substitués suivants : -(alkyle en C₁-C₄), -(alcoxy en C₁-C₄), -(alcényle en C₁-C₄), -(alcénoxy en C₁-C₄), -(alcynyle en C₁-C₄), -(alcynoxy en C₁-C₄), -(cycloalkyle en C₁-C₄) ou un hétérocycle, ayant de 1 à 3 atomes de carbone et 1 à 2 hétéroatomes ;

(iii) R² et R³ sont chacun indépendamment choisis parmi -H, -F, -Cl, -Br, -I, -CN, -OH (facultativement protégé avec un groupe protecteur d'hydroxyle), et -CH₂NH₂ (facultativement protégé avec un groupe protecteur d'azote) ;

(iv) R² et R³ sont chacun indépendamment choisis parmi les groupes facultativement substitués suivants : -(alkyle en C₁-C₄), -(alcoxy en C₁-C₄), -(alcényle en C₁-C₄), -(alcénoxy en C₁-C₄), -(alcynyle en C₁-C₄) et -(alcynoxy en C₁-C₄) ;

(v) R² et R³ facultativement substitué par un ou plusieurs groupes -OH ; ou

(vi) un ou plusieurs quelconques de R², R³ ou R⁵ est un -OH ou un groupe -OH protégé avec un groupe protecteur d'hydroxyle, facultativement où le groupe hydroxyle protecteur est choisi dans le groupe constitué

de benzyle, acétyle, benzoyle, tert-butyloxycarbonyl, silyle, tert- butyldiphénylsilyle, triméthylsilyle, para-méthoxybényle, benzyldine, diméthylacétal, et méthoxyméthyle, par exemple où le groupe silyle est -Si-(R^a) (R^b) (R^a) et R^a, R^b, et R^c sont chacun indépendamment choisis dans le groupe constitué de groupes alkyle en C₁-C₆, phényle, acétyle, et benzyle.

5

14. Procédé de la revendication 12 ou la revendication 13 dans lequel :

(i) R¹ est -F, R⁴ est -H, et R⁵ est -H ;

(ii) R² est -F ; et/ou

10

(iii) X est C ou S.

15. Procédé de l'une quelconque des revendications 10 à 14 dans lequel :

15

(i) le procédé comprend en outre une ou plusieurs étapes choisies parmi : l'extraction du composé (1) ; l'isolement du composé (1) ; la purification du composé (1) ; et la cristallisation du composé (1) ; et/ou

(ii) le polypeptide comprend une séquence d'acides aminés choisie dans le groupe constitué de SEQ ID NO: 168, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, et 166.

20

25

30

35

40

45

50

55

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

- US 6207822 B [0003]
- WO 200617257A2 A [0004]
- WO 2008085300 A1 [0004]
- WO 2008039829 A2 [0004]
- US 5886171 A [0005]
- US 5738321 A [0005]
- WO 20050066120 A [0005]
- WO 2007030721 A [0005]
- WO 2007120824 A [0005]
- WO 2007119106 A [0005]
- WO 2007072088 A [0005]
- US 6133001 A [0006]
- WO 2000060107 A [0006]
- WO 2008151324 A1 [0007] [0133]
- US 20100062499 A1 [0008] [0132] [0133] [0196]
- WO 2010025085 A [0009]
- WO 2009036404 A [0009]
- US 5559030 A [0104]
- US 5700670 A [0104]
- US 5891685 A [0104]
- US 6399339 B [0104]
- US 6645746 B [0104]
- WO 0140450 A [0104]
- US 20060286646 A [0104]
- US 6800477 B [0104]
- WO 2005054491 A [0104]
- US 20080318295 A1 [0132] [0133] [0196]
- US 20090093031 A1 [0132] [0133] [0196]
- US 20090155863 A1 [0132] [0133] [0196]
- US 20090162909 A1 [0132] [0133] [0196]
- US 20090191605 A1 [0132] [0133] [0196]
- US 20100055751 A1 [0132] [0133] [0196]
- WO 2010025238 A2 [0132] [0133] [0196]
- WO 2010025287 A2 [0132] [0133] [0196]
- WO 2009008908 A2 [0142]
- WO 9522625 A [0160]
- WO 970078 A [0160]
- WO 9735966 A [0160]
- WO 9827230 A [0160]
- WO 0042651 A [0160]
- WO 0175767 A [0160]
- US 6537746 B [0160]
- US 6117679 A [0160]
- US 6376246 B [0160]
- US 6586182 B [0160]
- US 20080220990 A1 [0160]
- US 20090312196 A1 [0160]
- US 5538867 A [0193]
- US 6495023 B [0193]
- WO 2000053731 A [0193]
- WO 2005018579 A [0194] [0195]
- US 20050095619 A [0194]
- US 20050153417 A [0194]
- WO 2008085300A1 A [0214]
- JP 2010083880 A [0216]
- WO 2010056788 A [0217]
- US 2010160282 A1 [0218]
- WO 2010100255 A [0218]
- US 2010152156 A1 [0219]
- US 20060195947 A [0221]

Non-patent literature cited in the description

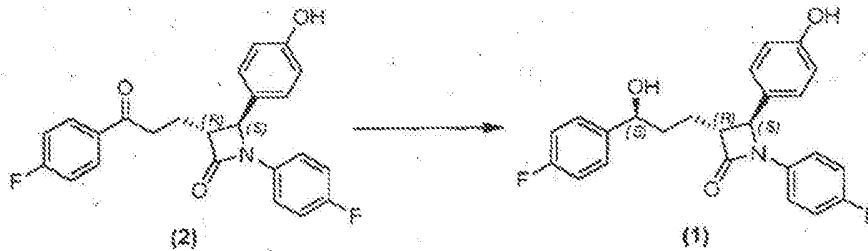
- **ALTSCHUL et al.** *J. Mol. Biol.*, 1990, vol. 215, 403-410 [0042]
- **ALTSCHUL et al.** *Nucleic Acids Res.*, 1977, 3389-3402 [0042]
- **HENIKOFF ; HENIKOFF.** *Proc Natl Acad Sci USA*, 1989, vol. 89, 10915 [0043]
- **SMITH ; WATERMAN.** *Adv. Appl. Math.*, 1981, vol. 2, 482 [0044]
- **NEEDLEMAN ; WUNSCH.** *J. Mol. Biol.*, 1970, vol. 48, 443 [0044]
- **PEARSON ; LIPMAN.** *Proc. Natl. Acad. Sci. USA*, 1988, vol. 85, 2444 [0044]
- *Current Protocols in Molecular Biology. Current Protocols*, 1995 [0044]
- **EISENBERG et al.** *J. Mol. Biol.*, 1984, vol. 179, 125-142 [0053] [0057]
- **ZHOU.** *J. Am. Chem. Soc.*, 1983, vol. 105, 5925-5926 [0104]
- **SANTANIELLO.** *J. Chem. Res. (S)*, 1984, 132-133 [0104]
- **KRAUS ; WALDMAN.** *Enzyme catalysis in organic synthesis.* VCH, 1995, vol. 1&2 [0105]
- **FABER, K.** *Biotransformations in organic chemistry.* Springer, 2000 [0105]
- **HUMMEL ; KULA.** *Eur. J. Biochem.*, 1989, vol. 184, 1-13 [0105]
- **FASMAN.** *CRC Practical Handbook of Biochemistry and Molecular Biology.* CRC Press, 1989, 3-70 [0137]

- **SAMBROOK et al.** *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, 2001 [0151]
- *Current Protocols in Molecular Biology*. Greene Pub. Associates, 1998 [0151]
- **STEMMER**. *Proc Natl Acad Sci USA*, 1994, vol. 91, 10747-10751 [0160]
- **ZHAO et al.** *Nat. Biotechnol.*, 1998, vol. 16, 258-261 [0161]
- **CALDWELL et al.** *PCR Methods Appl.*, 1994, vol. 3, S136-S140 [0161]
- **BLACK et al.** *Proc Natl Acad Sci USA*, 1996, vol. 93, 3525-3529 [0161]
- **LING et al.** *Anal. Biochem.*, 1997, vol. 254 (2), 157-78 [0161]
- **DALE et al.** *Methods Mol. Biol.*, 1996, vol. 57, 369-74 [0161]
- **SMITH**. *Ann. Rev. Genet.*, 1985, vol. 19, 423-462 [0161]
- **BOTSTEIN et al.** *Science*, 1985, vol. 229, 1193-1201 [0161]
- **CARTER**. Site-directed mutagenesis. *Biochem. J.*, 1986, vol. 237, 1-7 [0161]
- **KRAMER et al.** *Cell*, 1984, vol. 38, 879-887 [0161]
- **WELLS et al.** *Gene*, 1985, vol. 34, 315-323 [0161]
- **MINSHULL et al.** *Curr Opin Chem Biol*, 1989, vol. 3, 284-290 [0161]
- **CHRISTIANS et al.** *Nature Biotech*, 1999, vol. 17, 259-264 [0161]
- **CRAMERI et al.** *Nature*, 1998, vol. 391, 288-291 [0161]
- **CRAMERI et al.** *Nature Biotech*, 1997, vol. 15, 436-438 [0161]
- **ZHANG et al.** *Proc Natl Acad Sci USA*, 1997, vol. 94, 45-4, 4509 [0161]
- **CRAMERI et al.** *Nature Biotech*, 1996, vol. 14, 315-319 [0161]
- **STEMMER**. *Nature*, 1994, vol. 370, 389-391 [0161]
- **BEAUCAGE et al.** *Tet Lett*, 1981, vol. 22, 1859-69 [0163]
- **MATTHES et al.** *EMBO J.*, 1984, vol. 3, 801-05 [0163]
- **PUIGBÒ et al.** OPTIMIZER: a web server for optimizing the codon usage of DNA sequences. *Nucleic Acids Res.*, 16 April 2007, vol. 35, W126-31 [0221]

EZETIMIB-SZINTÉZIS BIOKATALIZÁTORAI

Szabadalmi igénypontok

1. Polipeptid, amely természetesen nem fordul elő, és amely képes a (2) vegyület szerinti ketont az (1) vegyület szerinti alkohollá alakítani diasztereómer feleslegben,



és amely polipeptid a SEQ ID NO:2 szerinti szekvenciával legalább 80% szekvenciaazonosságú aminosav-szekvenciát tartalmaz, és a következő jellemezők legalább egyikét tartalmazza:

- az X206-nak megfelelő pozícióban lévő csoport I,
- az X40-nak megfelelő pozícióban lévő csoport R,
- az X153-nak megfelelő pozícióban lévő csoport I vagy L,
- az X190-nak megfelelő pozícióban lévő csoport A vagy P,
- az X196-nak megfelelő pozícióban lévő csoport T és
- az X199-nak megfelelő pozícióban lévő csoport F vagy W.

2. Az 1. igénypont szerinti polipeptid, ahol

(i) az aminosav-szekvencia tartalmaz a következők közül választott jellemzőcsoportot is:

- az X93 pozícióban lévő csoport A és az X94 pozícióban lévő csoport T,
- az X93 pozícióban lévő csoport A és az X94 pozícióban lévő csoport S,
- az X93 pozícióban lévő csoport A és az X94 pozícióban lévő csoport P, vagy
- az X93 pozícióban lévő csoport I és az X94 pozícióban lévő csoport S,

(ii) az aminosav-szekvencia tartalmazza legalább a következő jellemzőket:

- az X40-nak megfelelő pozícióban lévő csoport R,
- az X93-nak megfelelő pozícióban lévő csoport A,
- az X94-nak megfelelő pozícióban lévő csoport T,
- az X153-nak megfelelő pozícióban lévő csoport I,
- az X190-nak megfelelő pozícióban lévő csoport A,
- az X196-nak megfelelő pozícióban lévő csoport T,
- az X199-nak megfelelő pozícióban lévő csoport F,
- az X203-nak megfelelő pozícióban lévő csoport G, és
- az X206-nak megfelelő pozícióban lévő csoport I,

(iii) az aminosav-szekvencia tartalmazza a következő jellemzőket:

- az X40-nak megfelelő pozícióban lévő csoport R,
- az X93-nak megfelelő pozícióban lévő csoport A,

az X94-nek megfelelő pozícióban lévő csoport S,
 az X96-nek megfelelő pozícióban lévő csoport N,
 az X153-nek megfelelő pozícióban lévő csoport L,
 az X190-nek megfelelő pozícióban lévő csoport A,
 az X196-nek megfelelő pozícióban lévő csoport T,
 az X199-nek megfelelő pozícióban lévő csoport F,
 az X202-nek megfelelő pozícióban lévő csoport G,
 az X203-nak megfelelő pozícióban lévő csoport G és
 az X206-nak megfelelő pozícióban lévő csoport I,

(iv) az aminosav-szekvencia tartalmazza a következő jellemzőket:

az X40-nek megfelelő pozícióban lévő csoport R,
 az X93-nak megfelelő pozícióban lévő csoport A,
 az X94-nek megfelelő pozícióban lévő csoport S,
 az X96-nak megfelelő pozícióban lévő csoport A, P vagy N,
 az X153-nak megfelelő pozícióban lévő csoport L,
 az X190-nek megfelelő pozícióban lévő csoport A,
 az X196-nak megfelelő pozícióban lévő csoport T,
 az X199-nek megfelelő pozícióban lévő csoport F,
 az X202-nek megfelelő pozícióban lévő csoport G,
 az X203-nak megfelelő pozícióban lévő csoport G és
 az X206-nak megfelelő pozícióban lévő csoport I,

ahol adott esetben az aminosav-szekvencia tartalmazza a következő jellemzőket is:

az X96-nak megfelelő pozícióban lévő csoport N és
 az X201-nek megfelelő pozícióban lévő csoport I vagy L,

(v) az aminosav-szekvencia tartalmazza legalább a következő jellemzőket:

az X40-nek megfelelő pozícióban lévő csoport R,
 az X93-nak megfelelő pozícióban lévő csoport A,
 az X96-nak megfelelő pozícióban lévő csoport P,
 az X153-nak megfelelő pozícióban lévő csoport L,
 az X190-nek megfelelő pozícióban lévő csoport A,
 az X196-nak megfelelő pozícióban lévő csoport T,
 az X199-nek megfelelő pozícióban lévő csoport F,
 az X202-nek megfelelő pozícióban lévő csoport G,
 az X203-nak megfelelő pozícióban lévő csoport G és
 az X206-nak megfelelő pozícióban lévő csoport I,

adott esetben ahol az aminosav-szekvencia a következő jellemzők közül egyet vagy többet is tartalmaz:

az X21-nek megfelelő pozícióban lévő csoport F,
 az X25-nek megfelelő pozícióban lévő csoport R,
 az X94-nak megfelelő pozícióban lévő csoport S,
 az X96-nak megfelelő pozícióban lévő csoport P,

az X108-nak megfelelő pozícióban lévő csoport K,
 az X127-nek megfelelő pozícióban lévő csoport K,
 az X147-nek megfelelő pozícióban lévő csoport M,
 az X153-nak megfelelő pozícióban lévő csoport L,
 az X205-nek megfelelő pozícióban lévő csoport V,
 az X211-nek megfelelő pozícióban lévő csoport K,
 az X221-nek megfelelő pozícióban lévő csoport D,

(vi) az aminosav-szekvencia tartalmazza legalább a következő jellemzőket:

az X21-nek megfelelő pozícióban lévő csoport F,
 az X25-nek megfelelő pozícióban lévő csoport R,
 az X40-nek megfelelő pozícióban lévő csoport R,
 az X93-nak megfelelő pozícióban lévő csoport A,
 az X94-nek megfelelő pozícióban lévő csoport S,
 az X96-nak megfelelő pozícióban lévő csoport P,
 az X108-nak megfelelő pozícióban lévő csoport K,
 az X127-nek megfelelő pozícióban lévő csoport K,
 az X147-nek megfelelő pozícióban lévő csoport M,
 az X153-nak megfelelő pozícióban lévő csoport L,
 az X190-nek megfelelő pozícióban lévő csoport A,
 az X196-nak megfelelő pozícióban lévő csoport T,
 az X199-nek megfelelő pozícióban lévő csoport F,
 az X202-nek megfelelő pozícióban lévő csoport G,
 az X203-nak megfelelő pozícióban lévő csoport G,
 az X206-nak megfelelő pozícióban lévő csoport I és
 az X221-nek megfelelő pozícióban lévő csoport D,

(vii) az aminosav-szekvencia tartalmazza legalább a következő jellemzőket:

az X40-nek megfelelő pozícióban lévő csoport R,
 az X93-nak megfelelő pozícióban lévő csoport I,
 az X94-nak megfelelő pozícióban lévő csoport S,
 az X96-nak megfelelő pozícióban lévő csoport P,
 az X153-nak megfelelő pozícióban lévő csoport L,
 az X190-nek megfelelő pozícióban lévő csoport A,
 az X196-nak megfelelő pozícióban lévő csoport T,
 az X199-nek megfelelő pozícióban lévő csoport F vagy W,
 az X201-nek megfelelő pozícióban lévő csoport A,
 az X202-nek megfelelő pozícióban lévő csoport G,
 az X203-nak megfelelő pozícióban lévő csoport G,
 az X206-nak megfelelő pozícióban lévő csoport I és
 az X211-nek megfelelő pozícióban lévő csoport K,

adott esetben ahol az aminosav-szekvencia a következő jellemzők közül egyet vagy többet is tartalmaz:

az X64-nek megfelelő pozícióban lévő csoport V,
 az X99-nek megfelelő pozícióban lévő csoport L,
 az X108-nak megfelelő pozícióban lévő csoport H,
 az X117-nek megfelelő pozícióban lévő csoport A vagy G,
 az X147-nek megfelelő pozícióban lévő csoport I vagy M,
 az X195-nek megfelelő pozícióban lévő csoport M,
 az X205-nek megfelelő pozícióban lévő csoport V,
 az X223-nak megfelelő pozícióban lévő csoport I,
 az X226-nak megfelelő pozícióban lévő csoport V,

(viii) az aminosav-szekvencia tartalmazza legalább a következő jellemzőket:

az X40-nek megfelelő pozícióban lévő csoport R,
 az X93-nak megfelelő pozícióban lévő csoport I,
 az X94-nak megfelelő pozícióban lévő csoport S,
 az X96-nak megfelelő pozícióban lévő csoport P,
 az X99-nek megfelelő pozícióban lévő csoport L,
 az X108-nak megfelelő pozícióban lévő csoport H,
 az X117-nek megfelelő pozícióban lévő csoport G,
 az X147-nek megfelelő pozícióban lévő csoport I,
 az X153-nak megfelelő pozícióban lévő csoport L,
 az X190-nek megfelelő pozícióban lévő csoport A,
 az X196-nak megfelelő pozícióban lévő csoport T,
 az X199-nek megfelelő pozícióban lévő csoport F,
 az X201-nek megfelelő pozícióban lévő csoport A,
 az X202-nek megfelelő pozícióban lévő csoport G,
 az X203-nak megfelelő pozícióban lévő csoport G,
 az X206-nak megfelelő pozícióban lévő csoport I,
 az X211-nek megfelelő pozícióban lévő csoport K és
 az X226-nak megfelelő pozícióban lévő csoport V, vagy

(ix) az aminosav-szekvencia tartalmazza a következő jellemzőket:

az X40-nek megfelelő pozícióban lévő csoport R,
 az X64-nek megfelelő pozícióban lévő csoport V,
 az X93-nak megfelelő pozícióban lévő csoport I,
 az X94-nak megfelelő pozícióban lévő csoport S,
 az X96-nak megfelelő pozícióban lévő csoport P,
 az X99-nek megfelelő pozícióban lévő csoport L,
 az X108-nak megfelelő pozícióban lévő csoport H,
 az X117-nek megfelelő pozícióban lévő csoport G,
 az X147-nek megfelelő pozícióban lévő csoport I,
 az X153-nak megfelelő pozícióban lévő csoport L,
 az X190-nek megfelelő pozícióban lévő csoport A,

az X196-nak megfelelő pozícióban lévő csoport T,
 az X199-nak megfelelő pozícióban lévő csoport W,
 az X201-nak megfelelő pozícióban lévő csoport A,
 az X202-nak megfelelő pozícióban lévő csoport G,
 az X203-nak megfelelő pozícióban lévő csoport G,
 az X206-nak megfelelő pozícióban lévő csoport I,
 az X211-nak megfelelő pozícióban lévő csoport K és
 az X226-nak megfelelő pozícióban lévő csoport V.

3. Az 1. igénypont vagy a 2. igénypont (i) pontja szerinti polipeptid, ahol az aminosav-szekvencia tartalmaz a következők közül választott jellemzőt vagy jellemzőcsoportot is:

az X203 pozícióban lévő csoport G,
 az X202 pozícióban lévő csoport G és az X203 pozícióban lévő csoport G, vagy
 az X201 pozícióban lévő csoport A, az X202 pozícióban lévő csoport G és az X203 pozícióban lévő csoport G.

4. Az 1-3. igénypontok bármelyike szerinti polipeptid, ahol az aminosav-szekvencia tartalmaz továbbá egyet vagy többet a következő jellemzők közül:

az X21-nak megfelelő pozícióban lévő csoport R vagy F,
 az X25-nak megfelelő pozícióban lévő csoport R, T vagy N
 az X40-nak megfelelő pozícióban lévő csoport R,
 az X64-nak megfelelő pozícióban lévő csoport V,
 az X93-nak megfelelő pozícióban lévő csoport A,
 az X94-nak megfelelő pozícióban lévő csoport T, S vagy P,
 az X95-nak megfelelő pozícióban lévő csoport V vagy M,
 az X96-nak megfelelő pozícióban lévő csoport V, G, A, N, S, P vagy T,
 az X99-nak megfelelő pozícióban lévő csoport L,
 az X108-nak megfelelő pozícióban lévő csoport H,
 az X117-nak megfelelő pozícióban lévő csoport A vagy G,
 az X127-nak megfelelő pozícióban lévő csoport K vagy Q,
 az X147-nak megfelelő pozícióban lévő csoport M vagy I,
 az X148-nak megfelelő pozícióban lévő csoport I,
 az X150-nak megfelelő pozícióban lévő csoport H vagy A,
 az X152-nak megfelelő pozícióban lévő csoport N vagy F,
 az X153-nak megfelelő pozícióban lévő csoport I vagy L, az X155-nak megfelelő pozícióban lévő csoport C,
 az X190-nak megfelelő pozícióban lévő csoport A,
 az X195-nak megfelelő pozícióban lévő csoport M,
 az X196-nak megfelelő pozícióban lévő csoport T, A, S, C vagy N,
 az X199-nak megfelelő pozícióban lévő csoport F vagy W,
 az X201-nak megfelelő pozícióban lévő csoport I, L vagy A,
 az X202-nak megfelelő pozícióban lévő csoport L, N, V vagy G,
 az X203-nak megfelelő pozícióban lévő csoport G,
 az X204-nak megfelelő pozícióban lévő csoport V vagy A,

az X205-nek megfelelő pozícióban lévő csoport V
 az X206-nak megfelelő pozícióban lévő csoport I,
 az X207-nek megfelelő pozícióban lévő csoport T, C, I vagy N,
 az X211-nek megfelelő pozícióban lévő csoport K,
 az X221-nek megfelelő pozícióban lévő csoport D,
 az X223-nak megfelelő pozícióban lévő csoport I,
 az X226-nak megfelelő pozícióban lévő csoport V.

5. Az 1. igénypont szerinti polipeptid, ahol

- (i) a polipeptid képes a (2) vegyületet legalább 99%-os diasztereomer feleslegben (1) vegyületté alakítani
- (ii) a polipeptid képes mintegy 100 g/l szubsztráatterhelés mellett a (2) vegyület legalább mintegy 91%-os, 92%-os, 93%-os, 94%-os, 95%-os, 96%-os, 97%-os, 98%-os, 99%-os vagy nagyobb arányú, (1) vegyületté iörténő átalakítására,
- (iii) a polipeptid képes a (2) vegyületet a SEQ ID NO: 2 szerinti polipeptid aktiválásához viszonyítva legalább kétszeresen, legalább tízszeresen, legalább 25-szörösen, legalább 40-szeresen vagy legalább 60-szorosan megnövekedett aktivitással (1) vegyületté alakítani,
- (iv) a polipeptid képes (2) vegyületet 24 óra alatt, 100 g/l szubsztráatterheléssel legalább 97%, 98, 99%, 99,1%, 99,2%, 99,3%, 99,4%, 99,5%, 99,6%, 99,7%, 99,8% vagy 99,9% vagy nagyobb (1) vegyület diasztereomer feleslegben (1) vegyületté alakítani, ahol adott esetben a reakciókörülmények magukban foglalnak továbbá mintegy 55-75% (v/v) IPA-koncentrációt, mintegy 0,03-0,5 g/l NADPH- vagy NADH-kofaktor-terhelést, és ahol a természetesen nem előforduló polipeptiden kívül nincs jelen kofaktort recikláló enzim, vagy
- (v) a polipeptid a következőkből álló csoportból választott aminosav-szekvenciát tartalmaz: SEQ ID NO: 168, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164 és 166.

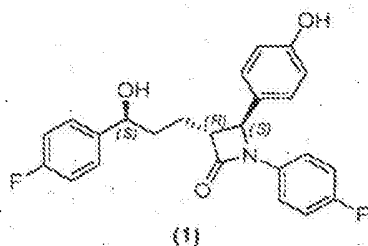
6. Polinukleotid, amely 1-5. igénypontok bármelyike szerinti polipeptidet kódol, és amely adott esetben a következőkből álló csoportból választott nukleotid-szekvenciát tartalmaz: SEQ ID NO: 167, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139, 141, 143, 145, 147, 149, 151, 153, 155, 157, 159, 161, 163 vagy 165.

7. Expressziós vektor, amely 6. igénypont szerinti polinukleotidot tartalmaz, és amely adott esetben kontrollszekvenciát tartalmaz, például ahol a kontrollszekvencia promotert tartalmaz.

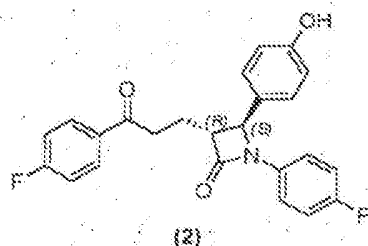
8. Gazdasejt, amely 6. igénypont szerinti polinukleotidot vagy 7. igénypont szerinti expressziós vektort tartalmaz, és amely adott esetben *E. coli*.

9. Eljárás 1-5. igénypontok bármelyike szerinti, természetesen nem előforduló polipeptid előállítására, amely magában foglalja 8. igénypont szerinti gazdasejt tenyésztését és a természetesen nem előforduló polipeptid izolálását a sejtéből.

10. Eljárás (1) vegyület vagy (1) vegyület analógjának előállítására diasztereomer feleslegben,



amely magában foglalja (2) vegyület vagy (2) vegyület analógjának érintkeztetését



1-5. igénypontok bármelyike szerinti polipeptiddel, NADPH vagy NADH kofaktorok jelenlétében, megfelelő reakciókörülmények között.

11. A 10. igénypont szerinti eljárás, ahol

- (i) a megfelelő reakciókörülmények magukban foglalják a következőket:
 - (1) szubsztrátterhelés mintegy 50-200 g/l (2) vegyülettel,
 - (2) mintegy 1,5-2,5 g/ polipeptid-koncentráció,
 - (3) mintegy 0,1-0,2 g/l NADPH-kofaktor-koncentráció,
 - (4) vizes puffer és mintegy 60-70% (v/v) IPA ko-szolvens oldat,
 - (5) mintegy 6,6-7,0 pH-érték és
 - (6) mintegy 28-30 °C hőmérséklet,
- (ii) a megfelelő reakciókörülmények magukban foglalnak legalább mintegy 20 g/l, mintegy 40 g/l, mintegy 50 g/l, mintegy 75 g/l, mintegy 100 g/l, mintegy 200 g/l, mintegy 250 g/l, mintegy 300 g/l vagy mintegy 400 g/l (2) vegyület terhelést
- (iii) a megfelelő reakciókörülmények magukban foglalnak mintegy 0,1-3,0 g/l, mintegy 0,5-2,75 g/l, mintegy 1,0-2,5 g/l, mintegy 1,5-2,5 g/l, mintegy 3 g/l, mintegy 2 g/l, mintegy 1,5 g/l, mintegy 1,0 g/l, mintegy 0,75 g/l vagy akár alacsonyabb módosított polipeptid-koncentrációt,
- (iv) a megfelelő reakciókörülmények magukban foglalnak mintegy 6,0-7,5, mintegy 6,25-7,25, mintegy 6,5-7,25, mintegy 6,6-7,25, mintegy 6,6-7,0, mintegy 6,75-7,25 vagy mintegy 6,75 pH-értéket,
- (v) a megfelelő reakciókörülmények magukban foglalnak mintegy 20-40°C, mintegy 23-37°C, mintegy 25-35°C, mintegy 26-32°C vagy mintegy 28-30°C hőmérsékletet,
- (vi) a megfelelő reakciókörülmények magukban foglalnak vizes pufferoldatot, szerves oldószert vagy ko-szolvens rendszert,
- (vii) a megfelelő reakciókörülmények magukban foglalnak mintegy 0,05-0,25 M TEA pufferoldatot,
- (viii) a megfelelő reakciókörülmények magukban foglalnak ko-szolvens rendszert, amely tartalmaz mintegy 30-70% (v/v) vizes pufferoldatot és mintegy 30-70% (v/v) szerves oldószert oldatot, adott esetben ahol a ko-szolvens rendszer tartalmaz vizes pufferoldatot, IPA-t és egyéb szerves oldószert, úgymint ahol:

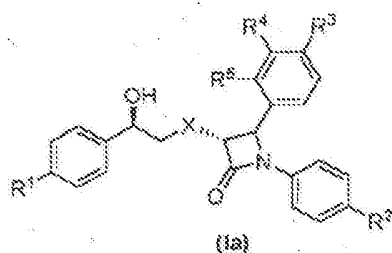
- (a) a ko-szolvens rendszer tartalmaz mintegy 45-55% (v/v) vizes pufferoldatot, mintegy 15-25% (v/v) IPA-t és mintegy 25-35% (v/v) toluolt, vagy
- (b) a ko-szolvens rendszer tartalmaz mintegy 50% (v/v) vizes pufferoldatot, mintegy 20% (v/v) IPA-t és mintegy 30% (v/v) toluolt,
- (ix) a megfelelő reakciókörülmények magukban foglalnak ko-szolvens rendszert, amely vizes pufferoldatot és IPA-t tartalmaz, ahol az IPA koncentrációja mintegy 25-75% (v/v), mintegy 35-75% (v/v), mintegy 45-75% (v/v), mintegy 55-75% (v/v), mintegy 60-70% (v/v), mintegy 62-68% (v/v), ahol adott esetben a ko-szolvens rendszer tartalmaz 0,1 M TEA-puffert és mintegy 60-70% (v/v) IPA-t, például ahol a ko-szolvens rendszer tartalmaz mintegy 35% (v/v) 0,1 M TEA-puffert és mintegy 65% (v/v) IPA-t,
- (x) a megfelelő reakciókörülmények magukban foglalnak mintegy 0,03-0,5 g/l, mintegy 0,05-0,3 g/l, mintegy 0,1-0,2 g/l, mintegy 0,5 g/l, mintegy 0,1 g/l vagy mintegy 0,2 g/l NADH- vagy NADPH-kofaktor-koncentrációt,
- (xi) az eljárás magában foglalja továbbá a következőt: NADP⁺ NADPH-vá alakítása kofaktor-recikláló rendszerrel, adott esetben ahol a kofaktor-recikláló rendszer ketoreduktáz és szekunder alkohol vagy glükóz-dehidrogenáz és glükóz közül választott, például ahol:
 - (a) a szekunder alkohol IPA, vagy
 - (b) a ketoreduktáz az 1-5. igénypontok bármelyike szerinti polipeptid,
- (xii) a megfelelő reakciókörülmények magukban foglalnak mintegy 55-75% (v/v) IPA-koncentrációt, mintegy 0,03-0,5 g/l NADPH- vagy NADH-kofaktor-terhelést, ahol az 1-5. igénypontok bármelyike szerinti polipeptiden kívül nincs jelen kofaktort recikláló enzim,
- (xiii) a megfelelő reakciókörülmények magukban foglalják a következőket:
 - (1) szubsztrátterhelés mintegy 100 g/l (2) vegyülettel,
 - (2) mintegy 2,0 g/l polipeptid-koncentráció,
 - (3) mintegy 0,1 g/l NADPH-kofaktor-koncentráció,
 - (4) 0,1 M TEA vizes puffere és mintegy 65% (v/v) IPA ko-szolvens oldata,
 - (5) mintegy 6,75 pH-érték és
 - (6) mintegy 30°C hőmérséklet,
- (xiv) a megfelelő reakciókörülmények magukban foglalják a következőket:
 - (1) szubsztrátterhelés mintegy 50-150 g/l (2) vegyülettel,
 - (2) mintegy 2,5-3,5 g/l polipeptid-koncentráció,
 - (3) mintegy 0,1-0,2 g/l NADPH-kofaktor-koncentráció,
 - (4) vizes pufferoldat és mintegy 15-25% (v/v) IPA és 25-35% (v/v) toluol ko-szolvens oldata,
 - (5) mintegy 6,6-7,0 pH-érték és
 - (6) mintegy 28-30°C hőmérséklet,
- (xv) a megfelelő reakciókörülmények magukban foglalják a következőket:
 - (1) szubsztrátterhelés mintegy 100 g/l (2) vegyülettel,
 - (2) mintegy 3,0 g/l polipeptid-koncentráció,
 - (3) mintegy 0,1 g/l NADPH-kofaktor-koncentráció,
 - (4) 0,1 M TEA vizes puffere és mintegy 20% (v/v) IPA és 30% (v/v) toluol ko-szolvens oldata,
 - (5) mintegy 6,6-7,0 pH-érték és

(6) mintegy 28-30°C hőmérséklet,

(xvi) a megfelelő reakciókörülmények magukban foglalják a következőt: szubsztráatterhelés mintegy 100 g/l (2) vegyülettel, ahol az eljárás 24 óra alatt (2) vegyület legalább mintegy 91%-os, 92%-os, 93%-os, 94%-os, 95%-os, 96%-os, 97%-os, 98%-os, 99%-os vagy nagyobb mértékű átalakulását eredményezi (1) vegyületté, vagy

(xvii) a megfelelő reakciókörülmények magukban foglalják a következőt: szubsztráatterhelés mintegy 100 g/l (2) vegyülettel, ahol az eljárás 24 óra alatt (1) vegyület legalább 97%, 98, 99%, 99,1%, 99,2%, 99,3%, 99,4%, 99,5%, 99,6%, 99,7%, 99,8% vagy 99,9% vagy nagyobb diasztereomer feleslegét eredményezi.

12. A 10. igénypont szerinti eljárás, ahol (1) vegyület analógja (1a) képletű vegyület,



ahol

X jelentése C vagy S,

R¹ jelentése -H, -F, -Cl, -Br és -I közül választott,

R² jelentése a következők közül választott: -H, -F, -Cl, -Br, -I, -CN, -OH (adott esetben hidroxil védőcsoporttal védett) -CH₂NH₂ (adott esetben nitrogén védőcsoporttal védett) és a következő, adott esetben szubsztituált csoportok: alkil, alkoxi, alkenil, alkenoxi, alkinil, alkinoxi, cikloalkil, aril, heteroaril vagy heterociklus,

R³ jelentése a következők közül választott: -H, -F, -Cl, -Br, -I, -CN, -OH (adott esetben hidroxil védőcsoporttal védett), -CH₂NH₂ (adott esetben nitrogén védőcsoporttal védett) és a következő, adott esetben szubsztituált csoportok: alkil, alkoxi, alkenil, alkenoxi, alkinil, alkinoxi, cikloalkil, aril, heteroaril vagy heterociklus,

R⁴ jelentése a következők közül választott: -H, -F, -Cl, -Br, -I, -CN, -OH (adott esetben hidroxil védőcsoporttal védett), és

R⁵ jelentése a következők közül választott: -H, -F, -Cl, -Br, -I, -CN, -OH (adott esetben hidroxil védőcsoporttal védett), -CH₂NH₂ (adott esetben nitrogén védőcsoporttal védett) és a következő, adott esetben szubsztituált csoportok: alkil, alkoxi, alkenil, alkenoxi, alkinil, alkinoxi, cikloalkil, aril, heteroaril vagy heterociklus

13. A 12. igénypont szerinti eljárás, ahol

(i) R² és R³ mindegyikének jelentése egymástól függetlenül a következők közül választott: -H, -F, -Cl, -Br, -I, -CN, -OH (adott esetben hidroxil védőcsoporttal védett), -CH₂NH₂ (adott esetben nitrogén védőcsoporttal védett) és a következő, adott esetben szubsztituált csoportok: -(C₁-C₆)alkil, -(C₁-C₆)alkoxi, -(C₁-C₆)alkenil, -(C₁-C₆)alkenoxi, -(C₁-C₆)alkinil, -(C₁-C₆)alkinoxi, -(C₁-C₆)cikloalkil és 1-4 szénatomot és 1-2 heteroatomot tartalmazó heterociklus,

(ii) R² és R³ mindegyikének jelentése egymástól függetlenül a következők közül választott: -H, -F, -Cl, -Br, -I, -CN, -OH (adott esetben hidroxil védőcsoporttal védett), -CH₂NH₂ (adott esetben nitrogén védőcsoporttal

- védett) és a következő, adott esetekben szubsztituált csoportok: $-(C_1-C_4)$ alkil, $-(C_1-C_4)$ alkoxi, $-(C_1-C_4)$ alkenil, $-(C_1-C_4)$ alkenoxi, $-(C_1-C_4)$ alkinil, $-(C_1-C_4)$ alkinoxi, $-(C_1-C_4)$ cikloalkil és 1-3 szénatomot és 1-2 heteroatomot tartalmazó heterociklus,
- (iii) R^2 és R^3 mindegyikének jelentése egymástól függetlenül a következők közül választott: -H, -F, -Cl, -Br, -I, -CN, -OH (adott esetekben hidroxil védőcsoporttal védett), és $-CH_2NH_2$ (adott esetben nitrogén védőcsoporttal védett),
- (iv) R^2 és R^3 mindegyikének jelentése egymástól függetlenül a következő, adott esetekben szubsztituált csoportok közül választott: $-(C_1-C_4)$ alkil, $-(C_1-C_4)$ alkoxi, $-(C_1-C_4)$ alkenil, $-(C_1-C_4)$ alkenoxi, $-(C_1-C_4)$ alkinil és $-(C_1-C_4)$ alkinoxi,
- (v) R^2 és R^3 adott esetekben egy vagy több -OH-csoporttal szubsztituált, vagy
- (vi) R^2 , R^3 vagy R^5 bármelyikének vagy R^2 , R^3 vagy R^5 közül többnek a jelentése -OH-csoport vagy hidroxil védőcsoporttal védett -OH-csoport, ahol adott esetekben a hidroxil védőcsoport a következőkből álló csoportból választott: benzil, acetyl, benzoyl, terc-butiloxikarbonil, szilil, terc-butildifenilszilil, trimetilszilil, para-metoxibenil, benzilidin, dimetilacetál és metoxi-metil, például ahol a szililcsoport $-Si-(R^a)(R^b)(R^c)$, és R^a , R^b és R^c mindegyikének jelentése egymástól függetlenül a következőkből álló csoportból választott: C_1-C_6 alkil-, fenil-, acetyl- és benzilcsoport.

14. A 12. vagy 13. igénypont szerinti eljárás, ahol:

- (i) R^1 jelentése -F, R^4 jelentése -H és R^5 jelentése -H;
- (ii) R^2 jelentése -F; és/vagy
- (iii) X jelentése C vagy S.

15. A 10-14. igénypontok bármelyike szerinti eljárás, ahol:

- (i) az eljárás magában foglal egy vagy több, a következők közül választott lépést: (1) vegyület extrahálása, (1) vegyület izolálása, (1) vegyület tisztítása és (1) vegyület kristályosítása; és/vagy
- (ii) a polipeptid a következőkből álló csoportból választott aminosav-szekvenciát tartalmaz: SEQ ID NO: 168, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164 és 166.