(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
10 August 2006 (10.08.2006)

PCT

(10) International Publication Number
WO 2006/082252 A2

(51) International Patent Classification:
C12P 7/42 (2006.01) ***C12N 1/20*** (2006.01)
C12P 13/06 (2006.01)

(21) International Application Number: PCT/EP2006/050722

(22) International Filing Date: 7 February 2006 (07.02.2006)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/650,136 7 February 2005 (07.02.2005) US

(71) Applicant *(for all designated States except US)*: **METABOLIC EXPLORER** [FR/FR]; Biopole Clermont-limagne, F-63360 Saint Beauzire (FR).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **FIGGE, Rainer** [DE/FR]; 32 rue Hippolyte Gomot, F-63200 Riom (FR). **LUX, Fabien** [FR/FR]; 20bis rue Amiral Gourbeyre, F-63200 Riom (FR). **RAYNAUD, Céline** [FR/FR]; 5 Impasse de la Source, F-63111 Dallet (FR). **SOUCAILLE, Philippe** [FR/FR]; Chant du Coucou, F-31450 Deyme (FR).

(74) Agent: CABINET REGIMBEAU-MARTIN, SCHRIMPF, WARCOIN, AHNER, TEXIER, LE FORESTIER, CALLON DE LAMARCK, COLLIN,; 129 Rue Servient, F-69326 Lyon Cedex 03 (FR).

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

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declaration under Rule 4.17:
— *of inventorship (Rule 4.17(iv))*

Published:
— *without international search report and to be republished upon receipt of that report*

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

(54) Title: METHOD FOR THE ENZYMATIC PRODUCTION OF ALPHA-KETOBUTYRATE

(57) Abstract: This invention relates to a process for the enzymatic production of α-ketobutyrate and its derivatives, in particular isoleucine. In the described process α-ketobutyrate is obtained from activated homoserine by γ-elimination. Homoserine is activated to either acylhomoserine or phosphohomoserine using homoserinetransacylase and homoserine kinase, respectively. Subsequently phosphohomoserine and/or acylhomoserine is/are transformed into phosphate or/and acylate, ammonia and α-ketobutyrate through the γ-eliminase activity of cystathionine- γ-synthase and/or acylhomoserine sulfliydrylase and/or phosphohomoserine sulfhydrylase. This reaction is favored by (i) reducing intracellular concentrations of H2S or cysteine, (ii) reducing the access of these two sulfur containing substrates to the enzyme cystathionine-γ- synthase and/or acylhomoserine sulfhydrylase and/or phosphohomoserine sulfliydrylase or (iii) by using cystathionine-γ-synthase and/or acylhomoserine sulfliydrylase and/or phosphohomoserine sulfhydrylase with high γ-elimination and low cystathionine-γ-synthase and/or low acylhomoserine sulfliydrylase and/or low phosphohomoserine sulfhydrylase activity. At least one of these measures is applied in microorganisms, in particular coryneform bacteria and enterobacteria, in which as a consequence α-ketobutyrate and/or its derivatives can be produced in high quantities.

WO 2006/082252 A2

Method for the enzymatic production of alpha-ketobutyrate

The present invention relates to a method for the enzymatic production of α-ketobutyrate, wherein α-ketobutyrate is obtained by γ-elimination of an activated homoserine with an appropriate enzyme in an appropriate medium, in particular where the appropriate enzymes are modified enzymes, preferably enzymes that when expressed from the plasmid pME101-thrA* in strains MG1655 Δ*met*BJ *met*A11 Δ*ilv*A Δ*tdc*ABCDEFG the ratio of isoleucine to methionine produced at least equal to two.

Background of the invention

L- amino acids and their derivatives are used in animal nutrition, human medicine and the pharmaceutical industry. Isoleucine, which is an essential amino acid, is principally used as a starting material for drug synthesis and in nutritional supplements.

Isoleucine that includes two asymmetric carbon atoms is difficult to synthesize chemically as the pure L-stereoisomer and therefore fermentative processes for the production of isoleucine have been developed (EP0745679B1, US5474918A1, EP0685555A1). These processes rely on the natural biosynthetic pathway in which the aspartate-derived homoserine is transformed into threonine that in turn can be converted by a set of five reactions to isoleucine (see Fig. 1). Conversion of aspartate to homoserine requires the enzyme aspartate semialdehyde dehydrogenase and aspartokinase/ homoserine dehydrogenase that as a fusion protein harbors both activities. Two Aspartokinases/homoserine dehydrogenases are present in *E. coli*, one is encoded by the gene *thr*A that is part of the threonine operon (encoding also homoserine kinase (*thr*B) and threonine synthase (*thr*C), which catalyze the biosynthesis of threonine) and the other is encoded by the gene *metL* that is part of the methionine regulon. The threonine operon is regulated by attenuation via threonine and isoleucine. In addition, the enzyme ThrA is feedback controlled by threonine, the end product of the pathway. Synthesis of isoleucine from threonine requires its deamination to α-ketobutyrate, a reaction catalyzed by two threonine deaminases IlvA and TdcB. IlvA, the major enzyme under aerobic conditions is feedback controlled by isoleucine. TdcB is preferentially expressed under anaerobic conditions. Subsequently α-ketobutyrate is transformed to isoleucine by a set of 4 additional reactions catalyzed by the enzymes acetohydroxyacid synthase (encoded by ilvB, ilvN, ilvG, ilvM, ilvL, ilvH) acetohydroxy acid isomeroreductase (ilvC), dihydroxyacid dehydratase (ilvD) and branched chained amino acid aminotransferase (ilvE), all of which are shared with valine biosynthesis. In *E. coli* three acetohydroxyacid synthases exist that are each preferentially feedback regulated by one of the three end-products isoleucine, valine and leucine (reviewed in Neidhardt, F. C. (Ed. in Chief), R. Curtiss III, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter, and H. E. Umbarger (eds). 1996. *Escherichia coli* and *Salmonella*: Cellular and Molecular Biology. American Society for Microbiology).

In addition to the synthesis via threonine deamination, α-ketobutyrate can also be derived from activated homoserine by γ-elimination, a side reaction of cystathionine-γ-synthase and/or acylhomoserine sulfhydrylase and/or phosphohomoserine sulfhydrylase. Cystathionine-γ-synthase and/or acylhomoserine sulfhydrylase and/or phosphohomoserine sulfhydrylase from different organisms will preferentially accept activated homoserines in the form of O-acetylhomoserine, O-succinylhomoserine and phosphohomoserine. In plants methionine and threonine are synthesized from phosphohomoserine, which is thus the branch-point between the two biosyntheses. In bacteria homoserine fulfils this role. To enter the methionine specific pathway, homoserine is either activated to acetyl- or succinyl-homoserine. In both plants and bacteria activated homoserine is the substrate of cystathionine-γ-synthase and/or acylhomoserine sulfhydrylase and/or phosphohomoserine sulfhydrylase that can catalyze one of the three reactions: (i) the synthesis of γ-cystathionine with cysteine, (ii) the sulfhydrylation to homocysteine with hydrogen sulfide or (iii) the γ-elimination to ammonia, phosphate and/or acylate and α-ketobutyrate in the absence of sulfur containing substrates. Cystathionine-γ-synthase and/or acylhomoserine sulfhydrylase and/or phosphohomoserine sulfhydrylase vary with respect to their ratios between the three reactions. Relatively low γ-eliminase activity has been observed in plant enzymes, whereas higher activity is found in the *E. coli* enzyme. γ-elimination activity also depends on the presence and the accessibility of the enzyme for the sulfur containing compounds.

Object of the invention is a process for the enzymatic production of isoleucine in which α-ketobutyrate is produced from activated homoserine. This can be accomplished by either reducing the amount of intracellular H_2S and/or cysteine, limiting access of these substrates to the active site of cystathionine-γ-synthase and/or acylhomoserine sulfhydrylase and/or phosphohomoserine sulfhydrylase or by using cystathionine-γ-synthase and/or acylhomoserine sulfhydrylase and/or phosphohomoserine sulfhydrylase with intrinsically high γ-elimination and low cystathionine-γ-synthase and/or acylhomoserine sulfhydrylase and/or phosphohomoserine sulfhydrylase activity.

Brief summary of the invention

The invention provides a new process for the production of α-ketobutyrate and its derivatives, in particular isoleucine by utilizing cystathionine-γ-synthase and/or acylhomoserine sulfhydrylase and/or phosphohomoserine sulfhydrylase, enzymes that normally are utilized in methionine biosynthesis. Enzymes that have high γ-elimination activity are preferred, since they will convert phospho and/or acylhomoserine into α-ketobutyrate. In addition, limiting conditions for the sulfur containing substrates H_2S and cysteine as well as reduced access to the active site of the acyl and/or phosphohomoserine cystathionine-γ-synthase and/or acylhomoserine sulfhydrylase and/or phosphohomoserine sulfhydrylase favor γ-elimination. Thus the invention provides a new pathway for the synthesis of isoleucine that deviates from the classical known pathway that relies on the synthesis of threonine and its subsequent deamination to α-ketobutyrate. As a consequence the invention makes the requirement for feedback resistant threonine deaminases obsolete.

Detailed description of the invention

The invention describes a process for the enzymatic production of α-ketobutyrate and its derivatives in particular isoleucine, in which α-ketobutyrate is derived from activated homoserine by γ-elimination. For this purpose the implicated enzymes have to be produced and active in an appropriate medium. An appropriate medium can be any solvent or mixture of solvents containing ions that allows the enzymes to be active in the production of α-ketobutyrate from activated homoserine. According to the invention activated homoserine can be phosphohomoserine or acylhomoserine, preferentially acetyl or succinylhomoserine.

The γ-elimination reaction can be performed using phosphohomoserine accepting cystathionine-γ–synthases and/or sulfhydrylases that are known in plants. Several amino acid positions are known to be relevant for the binding of the phosphogroup (Steegborn *et al.*, 1999, J. Mol. Biol. 290, 983-996). According to these conserved residues the inventors have deduced that cystathionine-γ–synthases and/or sulfhydrylases from the bacterial family Chloroflexaceae should also preferentially accept phosphohomoserine as a substrate. Recently it was shown that some archaebacterial species also use phosphohomoserine accepting cystathionine-γ–synthases and/or sulfhydrylases. Thus cystathionine-γ–synthases and/or sulfhydrylases may be derived from any of the organismal groups described above. Phosphohomoserine is obtained by phosphorylation of homoserine using homoserine kinase, which is preferentially homoserine kinase from *E. coli* or *Corynebacterium glutamicum*.

In another embodiment of the invention activated homoserine is acylhomoserine, preferentially acetylhomoserine. In this case the γ-elimination reaction is catalyzed by acetylhomoserine sulfhydrylases and/or cystathionine-γ-synthases that are preferentially derived from *Corynebacterium glutamicum*, *Bacillus subtilis* or *Saccharomyces cerevisiae*.

Acetylhomoserine is obtained by acetylation of homoserine, a reaction catalyzed by homoserine acetyltransferase. Preferentially homoserine transacetylases are employed that are feedback resistant to methionine and S-adenosylmethionine. In a preferred embodiment of the invention homoserine transacetylases are obtained from *Corynebacterium glutamicum, Bacillus subtilis* or *Saccharomyces cerevisiae* or a spirochaete.

In another application of the invention activated homoserine is succinylhomoserine. Succinylhomoserine is converted into succinate, ammonia and α-ketobutyrate by the γ-elimination reaction catalyzed by succinylhomoserine sulfhydrylase and/or cystathionine-γ-synthase. Preferentially succinylhomoserine sulfhydrylase and/or cystathionine-γ-synthase is obtained from enterobacteriaceae, specifically from *E. coli.* Homoserine is activated to succinylhomoserine by homoserine succinyltransferase. Homoserine succinyltransferases that are feedback resistant to methionine and S-adenosylmethionine are preferred and have been described in patent application PCT IB2004/001901. Homoserine succinyltransferases are derived from Enterobacteriaceae, preferentially from *E. coli.*

In the description of the present invention, enzymes are identified by their specific activities. This definition thus includes all polypeptides that have the defined specific activity also present in other organisms, more particularly in other microorganisms. Often enzymes with similar activities can be identified by their grouping to certain families defined as PFAM or COG.

PFAM (protein families database of alignments and hidden Markov models; http://www.sanger.ac.uk/Software/Pfam/) represents a large collection of protein sequence alignments. Each PFAM makes it possible to visualize multiple alignments, see protein domains, evaluate distribution among organisms, gain access to other databases, and visualize known protein structures.

COGs (clusters of orthologous groups of proteins; http://www.ncbi.nlm.nih.gov/COG/) are obtained by comparing protein sequences from 43 fully sequenced genomes representing 30 major phylogenic lines. Each COG is defined from at least three lines, which permits the identification of former conserved domains.

The means of identifying homologous sequences and their percentage homologies are well known to those skilled in the art, and include in particular the BLAST programs, which can be used from the website http://www.ncbi.nlm.nih.gov/BLAST/ with the default parameters indicated on that website. The sequences obtained can then be exploited (e.g., aligned) using, for example, the programs CLUSTALW (http://www.ebi.ac.uk/clustalw/) or MULTALIN (http://prodes.toulouse.inra.fr/multalin/cgi-bin/multalin.pl), with the default parameters indicated on those websites.

Using the references given on GenBank for known genes, those skilled in the art are able to determine the equivalent genes in other organisms, bacterial strains, yeasts, fungi, mammals, plants, etc. This routine work is advantageously done using consensus sequences that can be determined by carrying out sequence alignments with genes derived from other microorganisms, and designing degenerate probes to clone the corresponding gene in another organism. These routine methods of molecular biology are well known to those skilled in the art, and are described, for example, in Sambrook *et al.* (1989 Molecular Cloning: a Laboratory Manual. 2nd ed. Cold Spring Harbor Lab., Cold Spring Harbor, New York.).

In addition, the invention also related to enzymes that have been modified, especially optimized for the desired activity. In the case of the invention this optimization concerns specifically enzymes that activate homoserine, in particular the reduction of feedback by end-products can increase their activity. Equally important is the optimization or an increase in γ-eliminase activity while at the same time the cystathionine-γ-synthase and/or acylhomoserine sulfhydrylase and/or phosphohomoserine sulfhydrylase activities are reduced. This optimization can be related to the amount of isoleucine versus methionine produced in a strain in which the enzymes normally producing isoleucine are lacking (*ilv*A, *tdc*B); an example is the strain MG1655 Δ*metBJ* *met*A11 Δ*ilv*A Δ*tdc*ABCDEFG in which the meB alleles can be introduced on a low copy plasmid (Details for these experiments are given in example 1). In a preferred embodiment of the invention these strains produce a ratio of isoleucine to methionine that is at least of a factor of two or higher. The wildtype MetB allele could be shown to produce only a ratio of 1.4. Modified enzymes may be obtained by directed evolution, preferentially coupled with enzyme modeling, site directed mutagenesis or *in vivo* or *in vitro* evolution of the enzyme as described in (Directed Enzyme Evolution Screening and Selection Methods and Directed Enzyme Evolution Library Creation Methods and Protocols, 2003, eds Arnold, F. and Georgiou G., Humana Press).

The inventors have identified mutated versions of cystathionine-γ-synthase and/or acylhomoserine sulfhydrylase and/or phosphohomoserine sulfhydrylase with an increased ratio between *in vivo* γ-elimination activity and cystathionine-γ-synthase and/or acylhomoserine sulfhydrylase and/or phosphohomoserine sulfhydrylase. These mutated MetB enzymes are also object of the invention. Mutated MetB enzymes have preferentially at least one mutation in the following three conserved regions or combinations thereof.

All references to amino acid positions are made based on the cystathionine-γ-synthase/succinylhomoserine sulfhydrylase encoded by the *metB* gene of *E. coli* (SEQ ID NO 1).

>E. coli |EG10582|MetB: 386 aa - Cystathionine gamma-synthase

```
MTRKQATIAV RSGLNDDEQY GCVVPPIHLS STYNFTGFNE PRAHDYSRRG
NPTRDVVQRA LAELEGGAGA VLTNTGMSAI HLVTTVFLKP GDLLVAPHDC
YGGSYRLFDS LAKRGCYRVL FVDQGDEQAL RAALAEKPKL VLVESPSNPL
LRVVDIAKIC HLAREVGAVS VVDNTFLSPA LQNPLALGAD LVLHSCTKYL
NGHSDVVAGV VIAKDPDVVT ELAWWANNIG VTGGAFDSYL LLRGLRTLVP
RMELAQRNAQ AIVKYLQTQP LVKKLYHPSL PENQGHEIAA RQQKGFGAML
SFELDGDEQT LRRFLGGLSL FTLAESLGGV ESLISHAATM THAGMAPEAR
AAAGISETLL RISTGIEDGE DLIADLENGF RAANKG
```

The relative positions of corresponding conserved regions in other cystathionine-γ-synthase or acylhomoserine sulfhydrylase or phosphohomoserine sulfhydrylase from different organisms can be found by a person skilled in the art by simple sequence alignment as represented in figure 2 with the enzymes listed below:

>gi|2852454|dbj|BAA24699.1| cystathionine gamma-synthase, *Arabidopsis thaliana*
>gi|8439541|gb|AAF74981.1|AF082891_1 cystathionine gamma-synthase isoform 1, *Solanum tuberosum*
>gi|2198853|gb|AAB61348.1| cystathionine gamma-synthase, *Zea mays*
>gi|4322948|gb|AAD16143.1| cystathionine gamma-synthase precursor, *Nicotiana tabacum*
>gi|305042|gb|AAB03071.1| cystathionine gamma-synthase, *Escherichia coli*
>gi|21323418|dbj|BAB98046.1| O-acetylhomoserine sulfhydrylase, *Corynebacterium glutamicum* ATCC 13032
>gi|4191271|emb|CAA09983.1| O-succinylhomoserine sulfhydrylase, *Rhizobium etli*
>gi|53798754|ref|ZP_00020132.2| COG0626: Cystathionine beta-lyases/cystathionine gamma-synthases, *Chloroflexus aurantiacus*
>gi|14601268|ref|NP_147803.1| cystathionine gamma-lyase *Aeropyrum pernix* K1
>gi|26988726|ref|NP_744151.1| O-succinylhomoserine sulfhydrylase, *Pseudomonas putida* KT2440
>gi|15598303|ref|NP_251797.1| o-succinylhomoserine sulfhydrylase, *Pseudomonas aeruginosa* PAO1
>gi|53803333|ref|YP_114900.1| O-succinylhomoserine sulfhydrylase, *Methylococcus capsulatus* str. Bath
>gi|4191271|emb|CAA09983.1| O-succinylhomoserine sulfhydrylase, *Rhizobium etli*
>gi|41324877|emb|CAF19359.1| O-Acetylhomoserine (Thiol)-Lyase, *Corynebacterium glutamicum* ATCC 13032

>gi|38233237|ref|NP_939004.1| Putative methionine biosynthesis-related protein, *Corynebacterium diphtheriae* NCTC 13129

>gi|41409555|ref|NP_962391.1| MetB, *Mycobacterium avium* subsp. *paratuberculosis* str. k10

>gi|54025566|ref|YP_119808.1| putative O-acetylhomoserine sulfhydrylase, *Nocardia farcinica* IFM 10152

>gi|15896038|ref|NP_349387.1| O-acetylhomoserine sulfhydrylase, *Clostridium acetobutylicum* ATCC 824

>gi|45657715|ref|YP_001801.1| O-acetylhomoserine (thiol) lyase, *Leptospira interrogans* serovar Copenhageni str. Fiocruz L1-130

>gi|48866095|ref|ZP_00319952.1| COG2873: O-acetylhomoserine sulfhydrylase, *Oenococcus oeni* PSU-1

>gi|23024881|ref|ZP_00064071.1| COG2873: O-acetylhomoserine sulfhydrylase, *Leuconostoc mesenteroides* subsp. *mesenteroides* ATCC 8293

>gi|33865385|ref|NP_896944.1| putative O-Acetyl homoserine sulfhydrylase *Synechococcus* sp. WH 8102

>gi|39937649|ref|NP_949925.1| homocysteine synthase, *Rhodopseudomonas palustris* CGA009

>gi|33240250|ref|NP_875192.1| O-acetylhomoserine sulfhydrylase, *Prochlorococcus marinus* subsp. marinus str. CCMP1375]

>gi|20091539|ref|NP_617614.1| O-acetylhomoserine (thiol)-lyase, *Methanosarcina acetivorans* C2A

>gi|48838423|ref|ZP_00295367.1| COG2873: O-acetylhomoserine sulfhydrylase, *Methanosarcina barkeri* str. fusaro

In a preferred embodiment of the invention MetB has one or several amino acid changes in conserved region 1 comprising the following amino acids:

X1-X2-X3-**Y**-X4-**R**-X5-X6-**N**-**P**-**T**

in which

X1 represents S, R, G or is missing

X2 represents F, E, N, H, Y, P

X3 represents E, I, V, D, R

X4 represents G, A, S, T

X5 represents Y, F, R, S, T, L, I

X6 represents G, T, A, S, M

In another embodiment MetB has one or several amino acid changes in conserved region 2 comprising the following sequence:

X1-X2-X3-X4-X5-**G**-X6-X7-X8
In which
X1 represents I, H, L, N, R
X2 represents A, S, T, G, L, V,
X3 represents P, N, T, G, E, A, V
X4 represents S, N
X5 represents F, L, I, V
X6 represents G, D
X7 represents C, V, S, T, A,
X8 represents E, K, R

In another embodiment MetB has one or several amino acid changes in conserved region 3 comprising the following sequence:

X1-X2-**V/I**-X3-X4-**P/A**-X5-X6-X7-X8
In which
X1 represents S, T
X2 represents I, L, T
X3 represents D, E, T, A, S, C, V, I
X4 represents Q, H, V, I
X5 represents A, G, S, K
X6 represents I, T, S, R, V, L
X7 represents M, T
X8 represents S, T

Modified cystathionine-γ-synthase and/or acylhomoserine sulfhydrylase and/or phosphohomoserine sulfhydrylase comprises at least one of the following mutations and combinations thereof.

a valine at position X3 of conserved region 1
a leucine at position X5 of conserved region 1
a leucine or asparagine at position X1 of conserved region 2
an alanine, threonine or valine at position X3 of conserved region 2
an asparagine at position X4 of conserved region 2
a proline at the conserved position P/A
an aspartate at position X6 of conserved region 2
a lysine or arginine at position X8 of conserved region 2
a threonine at position X7 of conserved region 3.

The *met*B/Y/Z genes encoding modified cystathionine-γ-synthase and/or acylhomoserine sulfhydrylase and/or phosphohomoserine sulfhydrylase may be encoded chromosomally or extrachromosomally. Chromosomally there may be one or several copies on the genome that can be introduced by methods of recombination known to the expert in the field. Extrachromosomally the genes may be carried by different types of plasmids that differ with respect to their origin of replication and thus their copy number in the cell. They may be present as 1-5 copies, ca 20 or up to 500 copies corresponding to low copy number plasmids with tight replication (pSC101, RK2), low copy number plasmids (pACYC, pRSF1010) or high copy number plasmids (pSK bluescript II)

The *met*B gene may be expressed using promoters with different strength that need or need not to be induced by inducer molecules. Examples are the promoter Ptrc, Ptac, Plac, the lambda promoter cI or other promoters known to the expert in the field.

MetB expression may be boosted or reduced by elements stabilizing or destabilizing the corresponding messenger RNA (Carrier and Keasling (1998) Biotechnol. Prog. 15, 58-64) or the protein (e.g. GST tags, Amersham Biosciences)

The method as claimed by the invention also includes the conduction of the above described enzymatic reaction in cells comprising genes encoding the required enzymes to produce α-ketobutyrate from activated homoserine. In addition, these cells may express genes that are required for the subsequent conversion of α-ketobutyrate to isoleucine.

Cells used in the method according to the invention are eukaryotes or prokaryotes.

In preferred embodiments, cells are microorganisms, preferably selected among *S. cerevisiae, E. coli, or C. glutamicum.*

The enzymes with γ-eliminase activity expressed in said cells are preferably different from the native enzymes present in the same organism. They may be either mutated genes of the same species or native or mutated genes of other species.

According to a preferred embodiment of the present invention, the cell is transformed to introduce a gene coding for an enzyme with said γ-eliminase activity.

Such gene may be introduced by different means available to the man skilled in the art:

- modification of the native gene by homologous recombination to introduce mutations in the enzyme encoded by the said gene to enhance γ-eliminase activity;
- integrating into the genome of the microorganism a foreign gene coding for the selected enzyme known to have high γ-eliminase activity; said foreign gene being under control of regulatory elements functional in the host microorganism;

- introducing a plasmid comprising a foreign gene coding for the selected enzyme known to have high γ-eliminase activity under control of regulatory elements functional in the host microorganism.

When the gene is integrated into the genome of the microorganism, it may advantageously be introduced in a locus selected to replace the native gene.

Methods used to transform cells, and particularly microorganisms, including homologous recombination, are well known in the art.

Production of α-ketobutyrate or the derived isoleucine may be further increased by enhancing the expression of one or the following genes. Enhancing in this context means increasing the expression of a said gene and thus the activity of the corresponding enzyme:

Gene	*genbank entry*	*name*
*ace*K	g1790446	isocitrate dehydrogenase kinase/phosphatase
ppc	g1790393	phosphoenolpyruvate carboxylase
pps	g1787994	phosphoenolpyruvate synthase
*ilv*B	g1790104	acetohydroxy acid synthase I, large subunit
*ilv*N	g1790103	acetohydroxy acid synthase I, small subunit
*ilv*G	g1790202 g1790203	acetohydroxy acid synthase II, large subunit
*ilv*M	g1790204	acetohydroxy acid synthase II, small subunit
*ilv*I	g1786265	acetohydroxy acid synthase III, large subunit
*ilv*H	g1786266	acetohydroxy acid synthase III, small subunit
thrA	g1786183	homoserine dehydrogenase/ aspartokinase
pyc	U51439	pyruvate carboxylase or its homologs
*met*L	g1790376	homoserine dehydrogenase/aspartokinase
*lys*C	g1790455	apartokinase
asd	g1789841	aspartate semialdehyde dehydrogenase
*asp*C	g1787159	aspartate aminotransferase
*thr*B	g1786184	Homoserine kinase

Production of α-ketobutyrate or the derived isoleucine may be further increased by decreasing or completely eliminating the expression of one or the following genes:

Gene	*genbank entry*	*name*
*asp*A	*g1790581*	aspartate ammonia lyase
pck	g1789807	phosphoenolpyruvate carboxykinase
*ack*A	g1788633	acetate kinase

pta	g1788635	phosphotransacetylase
acs	g1790505	acetate synthase
*ace*E	g1786304	pyruvate deydrogenase E1
*ace*F	g1786305	pyruvate deydrogenase E2
lpd	g1786307	pyruvate deydrogenase E3
*suc*C	g1786948	succinyl-CoA synthetase, beta subunit
sucD	g1786949	succinyl-CoA synthetase, alpha subunit
pykA	g1788160	pyruvate kinase II
pykF	g1787965	pyruvate kinase I
poxB	g1787096	pyruvate oxidase
*aro*F	g1788953	DAHP synthetase
*aro*G	g1786969	DAHP synthetase
*aro*H	g1787996	DAHP synthetase
*cys*A	g1788761	sulfate permease
*cys*U	g1788764	cysteine transport system
*cys*W	g1788762	membrans bound sulphate transport system
cysZ	g1788753	ORF upstream of *cys*K
*cys*N	g1789108	ATP sulfurylase
*cys*D	g1789109	sulfate adenylyltransferase
*cys*C	g1789107	adenylylsulfate kinase
*cys*H	g1789121	adenylylsulfate reductase
*cys*I	g1789122	sulfite reductase, alpha subunit
*cys*J	g1789123	sulfite reductase, beta subunit
*cys*E	g1790035	serine acetyltransferase
*cys*K	g1788754	cysteine synthase
*cys*M	g2367138	O-acetyl-sulfhydrolase
*ser*A	g1789279	phosphoglycerate dehydrogenase
*ser*B	g1790849	phosphoserine phosphatase
*ser*C	g1787136	phosphoserine aminotransferase
*gly*A	g1788902	serine hydroxymethyltransferase
*met*F	g1790377	5,10-Methylenetetrahydrofolate reductase
*met*B	g1790375	Cystathionine gamma-synthase
*met*C	g1789383	Cystathionine beta-lyase
*met*H	g1790450	B12-dependent homocysteine-N5-methyltetrahydrofolate transmethylase

*met*E	g2367304	Tetrahydropteroyltriglutamate methyltransferase
*met*F	g1790377	5,10-Methylenetetrahydrofolate reductase
*met*R	g1790262	Positive regulatory gene for *met*E and *met*H and autogenous regulation
*met*J	g1790373	repressor of the methionine operon
*ast*A	g1788043	Arginine succinyltransferase
*dap*A	g1788823	Dihydrodipicolinate synthase

The invention also concerns the process for the production of α-ketobutyrate or its derivatives in particular isoleucine. α-ketobutyrate or its derivatives are usually prepared by fermentation of the designed bacterial strain.

According to the invention, the terms 'culture' and 'fermentation' are used indifferently to denote the growth of a microorganism on an appropriate culture medium containing a simple carbon source.

According to the invention a simple carbon source is a source of carbon that can be used by those skilled in the art to obtain normal growth of a microorganism, in particular of a bacterium. In particular it can be an assimilable sugar such as glucose, galactose, sucrose, lactose or molasses, or by-products of these sugars. An especially preferred simple carbon source is glucose. Another preferred simple carbon source is sucrose.

Those skilled in the art are able to define the culture conditions for the microorganisms according to the invention. In particular the bacteria are fermented at a temperature between 20°C and 55°C, preferentially between 25°C and 40°C, and more specifically about 30°C for *C. glutamicum* and about 37°C for *E. coli*.

The fermentation is generally conducted in fermenters with an inorganic culture medium of known defined composition adapted to the bacteria used, containing at least one simple carbon source, and if necessary a co-substrate necessary for the production of the metabolite.

In particular, the inorganic culture medium for *E. coli* can be of identical or similar composition to an M9 medium (Anderson, 1946, *Proc. Natl. Acad. Sci. USA* **32**:120-128), an M63 medium (Miller, 1992; A Short Course in Bacterial Genetics: A Laboratory Manual and Handbook for *Escherichia coli* and Related Bacteria, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York) or a medium such as defined by Schaefer *et al.* (1999, *Anal. Biochem.* **270**: 88-96).

Analogously, the inorganic culture medium for *C. glutamicum* can be of identical or similar composition to BMCG medium (Liebl *et al.*, 1989, *Appl. Microbiol. Biotechnol.* **32**:

205-210) or to a medium such as that described by Riedel *et al.* (2001, *J. Mol. Microbiol. Biotechnol.* **3**: 573-583). The media can be supplemented to compensate for auxotrophies introduced by mutations.

After fermentation α-ketobutyrate or its derivatives are recovered and purified if necessary. The methods for the recovery and purification of α-ketobutyrate or its derivatives such as isoleucine in the culture media are well known to those skilled in the art.

Description of figures:

Fig. 1 Metabolic pathway used for isoleucine production. Four different pathways are indicated: the original pathway that starts with the production of threonine, three new pathways that utilize activated homoserines, i.e. phosphohomoserine, acetylhomoserine and succinylhomoserine.

Fig. 2 Alignment of cystathionine-γ-synthases and acylhomoserine sulfhydrylases and phosphohomoserine sulfhydrylases, as indicated on page 6-8. Highly conserved residues are indicated in red, conserved residues in blue

Example 1:

Construction of an isoleucine producing mutant

To produce isoleucine via activated homoserine, homoserine can be activated by homoserine-transsuccinylase to succinylhomoserine. Subsequently succinylhomoserine is cleaved by cystathionine-γ-synthase/sulfhydrylase to succinate, NH_3, and α-ketobutyrate. In *Escherichia coli* this reaction sequence requires the expression of the genes *met*A and *met*B, encoding homoserine-transsuccinylase and cystathionine-γ-synthase/sulfhydrylase. Both genes are repressed by the repressor protein MetJ and its corepressor S-adenosyl methionine (SAM). In addition, the homoserine transsuccinylase activity is feedback controlled by methionine and SAM. To bypass these repression mechanisms a Δ*met*J mutant was constructed and feedback resistant mutants of MetA (encoded by *met*A^{fbr}) were introduced into the Δ*met*J mutant. These constructions have been described in patent application PCT IB2004/001901.

Under aerobic conditions threonine is deaminated to α-ketobutyrate by the enzyme threonine deaminase encoded by the *ilv*A gene. To demonstrate that isoleucine can truly be produced via activated homoserine, the strain ΔmetBJ *met*A^{fbr} was constructed starting with the *met*A^{fbr} already described (metA*11 in this case). To inactivate the *met*B and *met*J gene the homologous recombination strategy described by Datsenko & Wanner (2000) was used. This strategy allowed the insertion of a chloramphenicol resistance cassette, while deleting most of the genes concerned. For this purpose the following oligonucleotides were used:

DmetJR (SEQ ID NO 2)
tgacgtaggcctgataagcgtagcgcatcaggcgattccactccgcgccgctcttttttgctttagtattcccacg
tctcTGTAGGCTGGAGCTGCTTCG

with

- a region (lower case) homologous to the sequence (4125596-4125675) of the gene *met*J (reference sequence on the website http://genolist.pasteur.fr/Colibri/),
- a region (upper case) for the amplification of the chloramphenicol resistance cassette (reference sequence in Datsenko, K.A. & Wanner, B.L., 2000, PNAS, 97: 6640-6645),

DmetJBF (SEQ ID NO 3)
tatgcagctgacgacctttcgcccctgcctgcgcaatcacactcatttttaccccttgtttgcagcccggaagcca
ttttCAGGCACCAGAGTAAACATT

with

- a region (lower case) homologous to the sequence (4127460-4127381) of the gene *met*B and a region (4126116-4126197) homologous to the promoter of *met*L
- a region (upper case) for the amplification of the chloramphenicol resistance cassette.

The oligonucleotides DmetJBR and DmetJBF were used to amplify the chloramphenicol resistance cassette from the plasmid pKD3. The PCR product obtained was then introduced by electroporation into the strain MG1655 *met*A$^{\text{fbr}}$ (pKD46), described in patent application PCT IB2004/001901, in which the expressed Red recombinase enzyme permited the homologous recombination. The chloramphenicol resistant transformants were then selected and the insertion of the resistance cassette was verified by a PCR analysis with the oligonucleotides MetJR and MetJF defined below. The strain retained was designated MG1655 Δ*met*JB::Cm *met*A$^{\text{fbr}}$

MetJR (SEQ ID NO 4): ggtacagaaaccagcaggctgaggatcagc (homologous to the sequence from 4125431 to 4125460).

MetLR (SEQ ID NO 5): aaataacacttcacatcagccagactactgccaccaaattt (homologous to the sequence from 4127500 to 4157460).

The chloramphenicol resistance cassette was then eliminated. The plasmid pCP20 carrying recombinase FLP acting at the FRT sites of the chloramphenicol resistance cassette was introduced into the recombinant strains by electroporation. After a series of cultures at 42°C, the loss of the chloramphenicol resistance cassette was verified by a PCR analysis with the same oligonucleotides as those used previously.

Subsequently the *ilv*A gene was deleted using the same strategy and the following oligonucleotides:

DilvAF (SEQ ID NO 6)

GgctgactcgcaacccctgtccggtgctccggaaggtgccgaatatttaagagcagtgctgcgcgcgccggtttacgaggTGTAG GCTGGAGCTGCTTCG

with

- a region (lower case) homologous to the sequence (3952954-3953033) of the gene *ilv*A (reference sequence on the website http://genolist.pasteur.fr/Colibri/),
- a region (upper case) for the amplification of the chloramphenicol resistance cassette (reference sequence in Datsenko, K.A. & Wanner, B.L., 2000, PNAS, 97: 6640-6645),

DilvAR (SEQ ID NO 7)

cctgaacgccgggttattggtttcgtcgtggcaatcgtagcccagctcattcagccgggtttcgaaatccggttcatggCATATGA ATATCCTCCTTAG

with

- a region (lower case) homologous to the sequence 3954478-3954400) of the gene *ilv*A
- a region (upper case) for the amplification of the chloramphenicol resistance cassette.

The oligonucleotides DilvAR and DilvAF were used to amplify the chloramphenicol resistance cassette from the plasmid pKD3. The PCR product obtained was then introduced by electroporation into the strains MG1655 ΔmetBJ *met*A^{fbr} (pKD46) and ΔmetJ *met*A^{fbr} (pKD46) in which the expressed Red recombinase enzyme permitted the homologous recombination. The chloramphenicol resistant transformants were then selected and the insertion of the resistance cassette was verified by a PCR analysis with the oligonucleotides ilvAR and ilvAF defined below. The resulting strains are ΔmetBJ *met*A^{fbr} (*met*A*11 in this case) ΔilvA and ΔmetJ *met*A^{fbr} (*met*A*11 in this case) ΔilvA.

ilvAR (3954693-3954670) (SEQ ID NO 8): gccccgaaccggtgcgtaaccgcg

ilvAF (3952775-3952795) (SEQ ID NO 9): ggtaagcgatgccgaactggc

In addition to IlvA the threonine dehydratase (TdcB) is known to catalyze the deamination of threonine to α-ketobutyrate under anaerobic or microaerobic conditions. To eliminate the possible contribution of this enzyme to α-ketobutyrate production the gene was deleted from the genome of the strains Δ*met*J *met*A^{fbr} Δ*ilv*A and Δ*met*BJ *met*A^{fbr} Δ*ilv*A. TdcB is part of the operon tdcABCDEFG that was deleted in a similar way as described for previous mutants using the four oligonucleotides described below. DtdcGR and DtdcAF were used to amplify the cassette and tdcGR and tdcGF for the verification.

DtdcGR (3255915-3255993) (SEQ ID NO 10)

gctgacagcaatgtcagccgcagaccactttaatggccagtcctccgcgtgatgtttcgcggtatttatcgttcatatcCATATGAA TATCCTCCTTAG

DtdcAF (3264726-3264648) (SEQ ID NO 11)

GgtaattaacgtaggtcgttatgagcactattcttcttccgaaaacgcagcacctggtagtctttcaggaagtcattagTGTAGGCTGGAGCTGCTTCG

tdcGR (3255616-3255640) (SEQ ID NO 12)

gcgtctgcaatgacgcctttattcg

tdcAF (3264922-3264899) (SEQ ID NO 13)

Cgccataaaatatggttatccccg

Into the resulting strains Δ*met*J *met*A$^{\mathrm{fbr}}$ Δ*ilv*A Δ*tdc*ABCDEFG and Δ*met*BJ *met*A$^{\mathrm{fbr}}$ Δ*ilv*A Δ*tdc*ABCDEFG the plasmid pSB1 (for construction see below) was introduced to boost the production of homoserine, which can be transformed into α-ketobutyrate via its activated form.

To further boost the production of homoserine a *thr*A* allele with reduced feed-back resistance to threonine was expressed from the plasmid pCL1920 (Lerner & Inouye, 1990, NAR 18, 15 p 4631) using the promoter P*trc*. For the construction of plasmid pSB1 (pME101-thrA*1) *thr*A was PCR amplified from genomic DNA using the following oligonucleotides:

*Bsp*H1thrA (SEQ ID NO 14): ttaTCATGAgagtgttgaagttcggcggtacatcagtggc

*Sma*1thrA (SEQ ID NO 15): ttaCCCGGGccgccgccccgagcacatcaaacccgacgc

The PCR amplified fragment was cut with the restriction enzymes *Bsp*HI and *Sma*I and cloned into the *Nco*I / *Sma*I sites of the vector pTRC99A (Stratagene). For the expression from a low copy vector the plasmid pME101 was constructed as follows. The plasmid pCL1920 was PCR amplified using the oligonucleotides PME101F and PME101R and the *Bst*Z17I-*Xmn*I fragment from the vector pTRC99A harboring the *lac*I gene and the Ptrc promoter was inserted into the amplified vector. The resulting vector and the vector harboring the *thr*A gene were restricted by *Apa*I and *Sma*I and the *thr*A containing fragment was cloned into the vector pME101. To relieve ThrA from feed-back inhibition the mutation F318S was introduced by site-directed mutagenesis (Stratagene) using the oligonucleotides ThrAF F318S for and ThrAR F318S, resulting in the vector pME101-thrA*1, called pSB1.

PME101F (SEQ ID NO 16): Ccgacagtaagacgggtaagcctg

PME101R (SEQ ID NO 17): Agcttagtaaagccctcgctag

ThrAF F318S (SmaI) (SEQ ID NO 18): Ccaatctgaataacatggcaatg**tcc**agcgtttctggcccggg

ThrAR F318S (SmaI) (SEQ ID NO 19): Cccgggccagaaacgct**gga**cattgccatgttattcagattgg

Recombinant plasmids were verified by DNA sequencing.

Example 2:

Fermentation of the isoleucine producing mutants

Strains were initially analyzed in small Erlenmeyer flask cultures using modified M9 medium (Anderson, 1946, *Proc. Natl. Acad. Sci. USA* **32**:120-128) that was supplemented with 5 g/l MOPS and 5 g/l glucose. Spectomycin was added if necessary at a concentration of 50 mg/l. An overnight culture was used to inoculate a 30 ml culture to an OD600 of 0.2. After the culture had reached an OD600 of 4.5 to 5, 1,25 ml of a 50% glucose solution and 0.75 ml of a 2 M MOPS (pH 6.9) were added and the culture was agitated for 1 hour.

Extracellular metabolites were analyzed during the batch phase. Amino acids were quantified by HPLC after OPA/Fmoc derivatization and other relevant metabolites were analyzed using GC-MS after silylation.

The strain MG1655 *met*A$^{\text{fbr}}$ Δ*met*BJ Δ*ilv*A pSB1 was grown in the presence of either 2 mM methionine or 2 mM methionine and 2 mM isoleucine. In the absence of isoleucine the strain failed to grow demonstrating that *met*B is required for the production of isoleucine if the threonine deaminase encoding gene *ilv*A is deleted.

The following results were obtained for the strains, MG1655 Δ*met*J *met*A$^{\text{fbr}}$ pSB1 and MG1655 Δ*met*J *met*A$^{\text{fbr}}$ Δ*ilv*A Δ*tdc*ABCDEFG pSB1 grown without the addition of amino acids. Clearly the deletion of ΔilvA ΔtdcABCDEFG does not reduce isoleucine production indicating that isoleucine is produced by γ-elimination catalyzed by the MetB enzyme.

Product / strain	MG1655 Δ*met*J *met*A$^{\text{fbr}}$ pSB1	MG1655 Δ*met*J *met*A$^{\text{fbr}}$ Δ*ilv*A Δ*tdc*ABCDEFG pSB1
αKB	ND	ND
KMV	0.14	0.18
ISO	0.29	0.27
THR	0.02	0.02

Table 1. Concentrations of extracellular metabolites (mmol/gDW) after batch fermentation of strains MG1655 Δ*met*J *met*A$^{\text{fbr}}$ pSB1 and and MG1655 Δ*met*J *met*A$^{\text{fbr}}$ Δ*ilv*A Δ*tdc*ABCDEFG pSB1. Abbreviations: αKB α–Ketobutyrate , KMV 2-keto 3-methyl valerate, ISO isoleucine, THR threonine; ND not detected.

Strains that produced substantial amounts of metabolites of interest were subsequently tested under production conditions in 300 ml fermentors (DASGIP) using a fed batch protocol.

For this purpose the fermentor was filled with 145 ml of modified minimal medium containing 10 µM IPTG and inoculated with 5 ml of preculture to an optical density (OD600nm) between 0.5 and 1.2.

The temperature of the culture was maintained constant at 37 °C and the pH was permanently adjusted to values between 6.5 and 8 using an NH_4OH solution. The agitation rate was maintained between 200 and 300 rpm during the batch phase and was increased to up to 1000 rpm at the end of the fed-batch phase. The concentration of dissolved oxygen was maintained at values between 30 and 40% saturation by using a gas controller. When the optical density reached a value between three and five the fed-batch (medium containing 10 µM IPTG) was started with an initial flow rate between 0.3 and 0.5 ml/h and a progressive increase up to flow rate values between 2.5 and 3.5 ml/h. At this point the flow rate was maintained constant for 24 to 48 hours. The media of the fed was based on minimal media containing glucose at concentrations between 300 and 500 g/l.

When the concentration of biomass reached values between 30 and 60 g/l the fermentation was stopped and the extracellular isoleucine concentration was determined using HPLC.

strain	Isoleucine titer
MG1655 Δ*met*J *met*A^{fbr} Δ*ilv*A Δ*tdc*ABCDEFG pSB1	65 mM

Table 2. Isoleucine titer of strain MG1655 Δ*met*J *met*A^{fbr} Δ*ilv*A Δ*tdc*ABCDEFG pSB1 after fed batch fermentation.

Example 3:

Construction of cystathionine-γ-synthase/ sulfhydrylase mutants with increased γ-elimination activity

To increase the activity of cystathionine-γ-synthase/ sulfhydrylase (MetB) for γ-elimination, several mutations were introduced into regions that are involved in the binding of the substrate cysteine.

To this end *Escherichia coli met*B was PCR-amplified from genomic DNA using the oligonucleotides MetBF and MetBR (numbers in parentheses correspond to positions on the *E. coli* genome). The PCR fragment was restricted by *Pst*I and *Hind*III and cloned into pUC18 into the same restriction sites.

MetBF (4125957-4125982) (SEQ ID NO 20)

Ttagacagaa**ctgcag**cgccgctccattcagccatgagatac

MetBR (4127500-4127469) (SEQ ID NO 21)
Cgtaacgccc**aagctt**aaataacacttcacatcagccagactactgcc

Subsequently mutations were introduced into the *Escherichia coli* cystathionine-γ-synthase/ sulfhydrylase that resulted in the following amino acid changes: E325L, E325V, E325T, E325W and E325P using the following oligonucleotides. Mutant E325A had been previously obtained by a selection procedure that is described elsewhere (patent FR2851255). A pair of oligonucleotides was used for the introduction of each mutation using site directed mutagenesis according to Stratagene's Quick change^{TM} site directed mutagenesis KIT. The mutations are indicated in the name of the oligonucleotide.

E325PF (SEQ ID NO 22)
Cgttgtttacgctggcgccgtcattagggggagtggaaag
E325PR (SEQ ID NO 23)
ctttccactccccctaatgacggcgccagcgtaaacaacg
E32LF (SEQ ID NO 24)
Cgttgtttacgctggcgctgtcattagggggagtggaaag
E325LR (SEQ ID NO 25)
Ctttccactccccctaatgacagcgccagcgtaaacaacg
E325VF (SEQ ID NO 26)
cgttgtttacgctggcggtgtcattagggggagtggaaag
E325VF (SEQ ID NO 27)
ctttccactccccctaatgacaccgccagcgtaaacaacg
E325TF (SEQ ID NO 28)
cgttgtttacgctggcgacctcattagggggagtggaaag
E325TF (SEQ ID NO 29)
ctttccactccccctaatgaggtcgccagcgtaaacaacg
E325WF (SEQ ID NO 30)
cgttgtttacgctggcgtggtcattagggggagtggaaag
E325WR (SEQ ID NO 31)
ctttccactccccctaatgaccacgccagcgtaaacaacg

The resulting modified *met*B genes were verified by sequencing. Clones were restricted with *Pst*I and *Hind*III and the *met*B containing fragments cloned into the same sites of vector pSB1. Then plasmids were transformed into the strain MG1655 Δ*met*BJ *met*A11 Δ*ilv*A Δ*tdc*ABCDEFG.

Example 4:

Fermentation of strains with mutant MetB enzymes that increase the ratio of isoleucine to methionine produced

The strains MG1655 Δ*met*BJ *met*A11 Δ*ilv*A Δ*tdc*ABCDEFG expressing the different metB alleles were analyzed for isoleucine production as described in example 2. The strains with the mutant metB alleles showed an increased ratio of isoleucine to methionine produced, indicating that γ-elimination activity of the corresponding enzymes has been increased while cystathionine-γ-synthase and sulfhydrylase activities remain equal or decrease.

Strain	metBwt	E325P	E325A	E325L	E325V	E325T	E325W
methionine	0.96	0.15	0.14	0.15	0.12	0.15	0.15
isoleucine	0.88	0.68	0.44	0.69	0.71	0.34	0.78

Table 3 Methionine and isoleucine production (mmol/g Dw) of strain DmetBJ metA*11 pME101-*thr*A*1-*met*Bxx, where *met*Bxx is either wildtype *met*B (*met*Bwt) or *met*B with the amino acid changes indicated.

Example 5

Fermentation under sulfur limiting conditions

α-ketobutyrate production can be further increased by fermenting strains with high γ-eliminase activity under sulfur limiting conditions. These can either be achieved by reducing the assimilation of inorganic sulfur through mutations in the *cys* operon or by fermenting the strains under sulfur limiting conditions. Under these conditions γ-elimination will be favored and higher isoleucine titers can be obtained.

CLAIMS

1) A method for the enzymatic production of α-ketobutyrate, wherein α-ketobutyrate is obtained by γ-elimination of an activated homoserine with an appropriate enzyme in an appropriate medium.

2) The method as claimed in claim 1, wherein activated homoserine is phosphohomoserine.

3) The method as claimed in claim 2, wherein the appropriate enzyme is cystathionine-γ-synthase and/or phosphohomoserine sulfhydrylase.

4) The method as claimed in claim 3, wherein the cystathionine-γ-synthase and/or phosphohomoserine sulfhydrylase is obtained from plants, Chloroflexaceae or archaea.

5) The method as claimed in one of claims 2 to 4, wherein phosphohomoserine is obtained by enzymatic transformation of homoserine with a homoserine kinase.

6) The method as claimed in claims 4, wherein the homoserine kinase is obtained from *E. coli* or *Corynebacterium glutamicum*.

7) The method as claimed in claim 1, wherein activated homoserine is acylhomoserine.

8) The method as claimed in claim 7, wherein the appropriate enzyme is acetylhomoserine sulfhydrylase and/or cystathionine-γ-synthase.

9) The method as claimed in claim 8, wherein the cystathionine-γ-synthase and/or acetylhomoserine sulfhydrylase is obtained from *Corynebacterium glutamicum*, *Bacillus subtilis* or *Saccharomyces cerevisiae*.

10) The method as claimed in one of claims 7 to 9, wherein acylhomoserine is obtained by enzymatic transformation of homoserine with homoserine transacetylase.

11) The method as claimed in claim 10, wherein homoserine transacetylase is feedback resistant to methionine and S-adenosyl-methionine.

12) The method as claimed in one of claims 10 or 11, wherein the homoserine transacetylase is obtained from *Corynebacterium glutamicum, Saccharomyces cerevisiae* or from a *spirochete.*

13) The method as claimed in claim 1, wherein activated homoserine is succinylhomoserine.

14) The method as claimed in claim 13, wherein the appropriate enzyme is succinylhomoserine sulfhydrylase and/or cystathionine-γ-synthase.

15) The method as claimed in claim 14, wherein the cystathionine-γ-synthase and/or succinylhomoserine sulfhydrylase is derived from *Enterobacteriaceae*, preferentially from *E. coli.*

16) The method as claimed in one of claims 13 to 15, wherein succinylhomoserine is obtained by enzymatic transformation of homoserine with homoserine transsuccinylase.

17) The method as claimed in claim 16, wherein homoserine transsuccinylase is feedback resistant to methionine and S-adenosyl-methionine.

18) The method as claimed in one of claims 16 or 17, wherein the homoserine transsuccinylase is obtained from *Enterobacteriaceae,* preferentially from *E. coli.*

19) The method as claimed in one of claims 1 to 18, wherein the appropriate enzymes are modified enzymes.

20) The method as claimed in claim 19, wherein modified enzymes expressed from the plasmid pME101-thrA* in strain MG1655 Δ*met*BJ *met*A11 Δ*ilv*A Δ*tdc*ABCDEFG have a ratio of isoleucine to methionine produced equal to two or higher.

21) The method as claimed in one of claims 19 or 20, wherein modified cystathionine-γ-synthase and/or acylhomoserine sulfhydrylase and/or phosphohomoserine sulfhydrylase comprise at least one amino acid mutation when compared to the wild-type sequence in one of the following conserved regions and combinations thereof:

conserved region 1 comprising the following amino acids:

X1-X2-X3-**Y**-X4-**R**-X5-X6-**N**-**P**-**T**

in which

X1 represents S, R, G or is missing

X2 represents F, E, N, H, Y, P

X3 represents E, I, V, D, R

X4 represents G, A, S, T

X5 represents Y, F, R, S, T, L, I

X6 represents G, T, A, S, M

Conserved region 2 comprising the following sequence:

X1-X2-X3-X4-X5-**G**-X6-X7-X8

In which

X1 represents I, H, L, N, R

X2 represents A, S, T, G, L, V,

X3 represents P, N, T, G, E, A, V

X4 represents S, N

X5 represents F, L, I, V

X6 represents G, D
X7 represents C, V, S, T, A,
X8 represents E, K, R

Conserved region 3 comprising the following sequence:
X1-X2-**V/I**-X3-X4-**P/A**-X5-X6-X7-X8
In which
X1 represents S, T
X2 represents I, L, T
X3 represents D, E, T, A, S, C, V, I
X4 represents Q, H, V, I
X5 represents A, G, S, K
X6 represents I, T, S, R, V, L
X7 represents M, T
X8 represents S, T.

22) The method as claimed in claim 21, wherein the modified cystathionine-γ-synthase and/or acylhomoserine sulfhydrylase and/or phosphohomoserine sulfhydrylase comprise at least one of the following mutations and combinations thereof:
a valine at position X3 of conserved region 1
a leucine at position X5 of conserved region 1
a leucine or asparagine at position X1 of conserved region 2
an alanine, threonine, valine, tryptophane, praline or leucine at position X3 of conserved region 2
an asparagine at position X4 of conserved region 2
an aspartate at position X6 of conserved region 2
a lysine or arginine at position X8 of conserved region 2
a proline at the conserved position P/A of conserved region 3
a threonine at position X7 of conserved region 3.

23) The method as claimed in one of claims 1 to 22, wherein α-ketobutyrate is further converted into one of its derivatives, in particular isoleucine.

24) The method as claimed in any one of claims 1 to 23, wherein the enzymatic reaction is conducted in cells comprising genes coding for the appropriate enzymes as described in claims 3, 4, 8, 9, 14, 15, 19, 20, 21 and 22 , eventually genes coding for enzymes as described in claims 5, 6, 10, 11, 12, 16, 17 and 18, and eventually genes encoding enzymes appropriate for the conversion of α-ketobutyrate into one of its derivatives, in particular isoleucine.

25) The method as claimed in claim 24, wherein at least one of the said genes is enhanced, in particular overexpressed.

26) The method as claimed in one of claims 24 or 25, wherein at least one of the following genes is enhanced, in particular overexpressed:
*ace*K, *ppc, pps, ilv*B, *ilv*N, *ilv*G, *ilv*M, *ilv*I, *ilv*H, *thr*A, *pyc*, *met*L, *lys*C, *asd*, *asp*C, *thr*B,

27) The method as claimed in one of claims 24 to 26, wherein one or several of the following genes are attenuated, in particular deleted:
*asp*A, *pck, ack*A, *pta, acs*, *ace*E, *ace*F, *lpd, suc*C, *sucD, pykA, pykF, poxB, aro*F, *aro*G, *aro*H, *cys*A, *cys*U, *cys*W, *cysZ*, *cys*N, *cys*D, *cys*C, *cys*H, *cys*I, *cys*J, *cys*E, *cys*K, *cys*M, *ser*A, *ser*B, *ser*C, *gly*A, *met*F, *met*B, *met*C, *met*H, *met*E, *met*F, *met*R, *met*J, *ast*A, *dap*A.

28) The method as claimed in one of claims 24 to 27, wherein the cell is a microorganism, preferably selected among *E. coli*, *C. glutamicum* or *S. cerevisiae*.

29) The method as described in claim 28 wherein the microorganism is used for the fermentative production of α-ketobutyrate or its derivatives.

30) The method as claimed in one of claims 1 to 29, wherein α-ketobutyrate or its derivatives, in particular isoleucine, are concentrated in the medium or in the cells of the microorganisms.

31) The method as claimed in one of claims 1 to 30, wherein α-ketobutyrate or its derivatives, in particular isoleucine, are isolated.

32) A cell as disclosed in one of claims 24 to 28.

Fig. 1

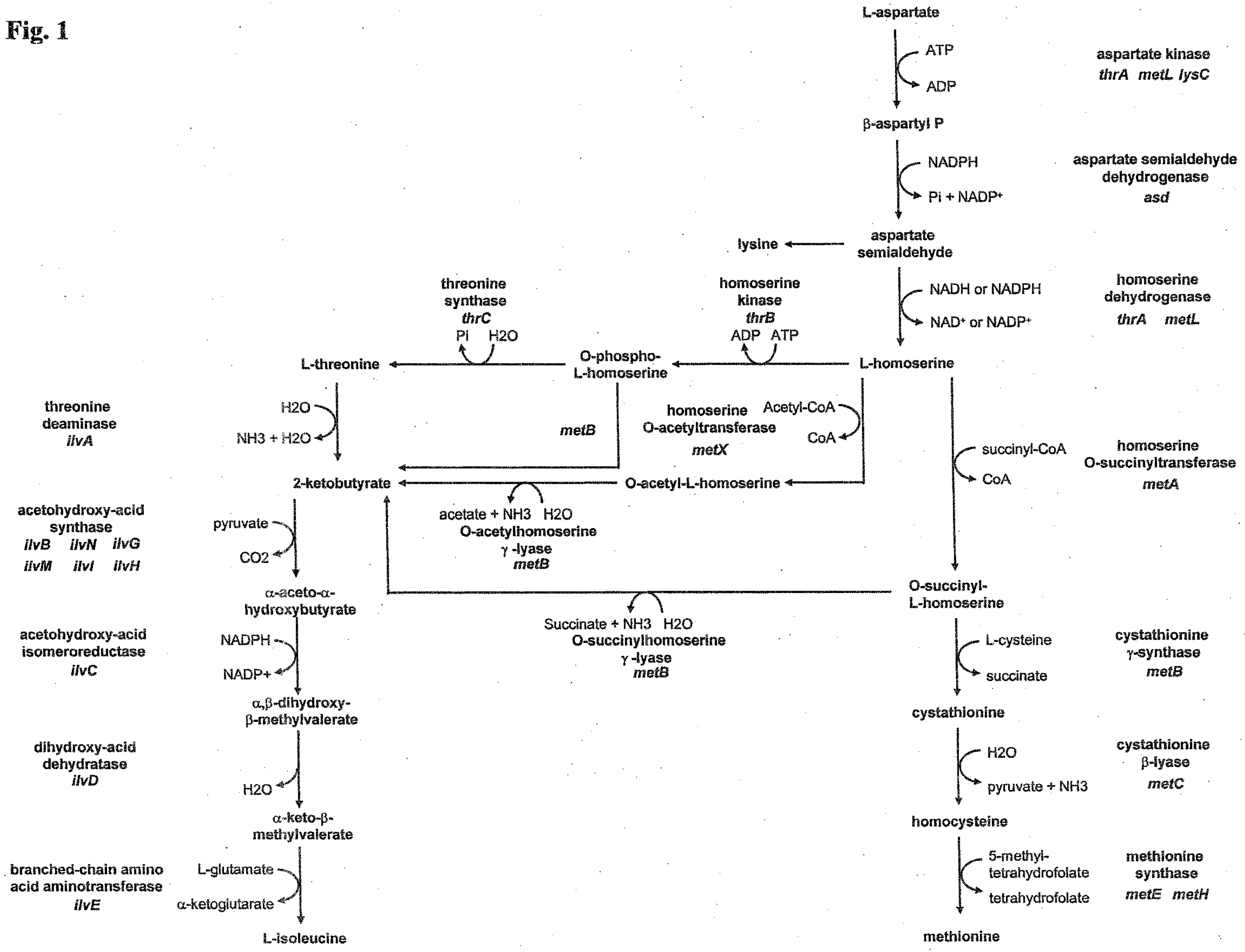
L-aspartate
ATP
ADP
aspartate kinase
thrA metL lysC
β-aspartyl P
NADPH
Pi + NADP+
aspartate semialdehyde dehydrogenase
asd
lysine
aspartate semialdehyde
NADH or NADPH
NAD+ or NADP+
homoserine dehydrogenase
thrA metL
threonine synthase
thrC
Pi H2O
homoserine kinase
thrB
ADP ATP
L-threonine
O-phospho-L-homoserine
L-homoserine
threonine deaminase
ilvA
H2O
NH3 + H2O
metB
homoserine O-acetyltransferase
metX
Acetyl-CoA
CoA
succinyl-CoA
CoA
homoserine O-succinyltransferase
metA
2-ketobutyrate
O-acetyl-L-homoserine
acetate + NH3 H2O
O-acetylhomoserine
γ -lyase
metB
acetohydroxy-acid synthase
ilvB ilvN ilvG
ilvM ilvI ilvH
pyruvate
CO2
α-aceto-α-hydroxybutyrate
O-succinyl-L-homoserine
Succinate + NH3 H2O
O-succinylhomoserine
γ -lyase
metB
acetohydroxy-acid isomeroreductase
ilvC
NADPH
NADP+
L-cysteine
succinate
cystathionine γ-synthase
metB
α,β-dihydroxy-β-methylvalerate
cystathionine
dihydroxy-acid dehydratase
ilvD
H2O
H2O
pyruvate + NH3
cystathionine β-lyase
metC
α-keto-β-methylvalerate
homocysteine
branched-chain amino acid aminotransferase
ilvE
L-glutamate
α-ketoglutarate
5-methyl-tetrahydrofolate
tetrahydrofolate
methionine synthase
metE metH
L-isoleucine
methionine

Fig. 2

```
gi|2852454  MAVSSFQCPT IFSSSSISGF QCRSDPDLVG SPVGGSSRRR VHASAGISSS FTGDAGLSSR ILRFPPNFVR QLSIKARRNC SNIGVAQIVA AKWSNNPSSA LPSAAAAAAT SSASAVSSAA SAAAASSAAA
gi|8439541  MAVSSY---- ---ARAFPSF ECRSEPDFSG SLPHPKAGVR FSGKYNSGSN RSQVYGLSSL IFRFPPNFVR QLSIKARRNC SNIGVAQVVA ASWSNNQAGP EFTPAANAVD SSASAAVTSI GITAGDEEVA
gi|4322948                                                                                                                    MAKAVDAA AAAAAAIAPV DTTVVNEDVA
gi|2198853                                       MATV SLTPQAVFST ESGGALASAT ILRFPPNFVR QLSTKARRNC SNIGVAQIVA AAWSDCPAAR PHLGGGGRRA RGVASSHAAA ASAAAAASAA
gi|5379875
gi|4191271
gi|2698872
gi|1559830
gi|5380333
gi|3050421
gi|5012317
gi|2889953
gi|2122333
gi|1567276
gi|1460126
gi|4132487
gi|3823323
gi|4140955
gi|5402556
gi|1589603
gi|4886609
gi|2302488
gi|4565771
gi|4883842                                                                                                                                         MKDKNRN
gi|3386538
gi|3324025
gi|3993764
  Consensus                                                                                                                   ........ .......... ..........
```

Fig. 2(to be continued)

```
gi|2852454  APVAAAPPVV LKSVDEEVVV AE-EGIREKI GSVQLTDSKH SFLSSDGSLT VHAGERLGRG IVTDAITTPV VNTSAYFFKK TAELIDFKEK RSVSFEYGRY GNPTTVVLED KISALEGAES ---TLVMASG
gi|8439541  VVENA----- -DCNDQNVQI TG-TGVKYA- ---------- SFLNSDGSVA IHAGERLGRG IVTDAITTPV VNTSAYFFNK TSDLIDFKEK RRASFEYGRY GNPTTVVLEE KISALEGAES ---TLIVASG
gi|4322948  LVENE----- -TCNDQNVQF DSLPSMKYA- ---------- SFLNSDGSVA IHAGERLGRG IVTDAITTPV VNTSAYFFNK TSELIDFKEK RRASFEYGRY GNPTTVVLEE KISALEGAES ---TLLMASG
gi|2198853  AEVSA----- --IPNAKVAQ PSAVVLAER- ---------- NLLGSDASLA VHAGERLGRR IATDAITTPV VNTSAYWFNN SQELIDFKEG RHASFEYGRY GNPTTEALEK KMSALEKAES ---TVFVASG
gi|5379875                                          MAI DQFTGPSTAA VHAGEP--RQ RAFDAITPPV VHSATYTFRD SAELIAFQSG DLEREEYGRY GNPTVRAVEA RLAALESPTA PAAALLCASG
gi|4191271                                            M TKTWRPATQL VHGGT---WR SEYGETSEAI YLTQGFVYDT SEPAEARFKG ETDGFIYARY GSPTNDMFEK RMCALEGAED ---ARATASG
gi|2698872                     MTD QWDAGRLDS- ---------D LEGVGFDTLA VRAGQ---NR TPEGEHSEAL FLTSSYVFRT AADAAARFAG ETPGNVYSRY TNPSVRAFEE RLAAMEGAEQ ---AVGTSTG
gi|1559830                     MTQ DWDAGRLDS- ---------D LEGAAFDTLA VRAGQ---RR TPEGEHGEAL FTTSSYVFRT AADAAARFAG EVPGNVYSRY TNPTVRTFEE RIAALEGAEQ ---AVATASG
gi|5380333                                                    METRA VRAGQ---RR TMEQEHAEPI FATSSYVFAS AAEAAERFAG KAAGNIYSRF TNPTVRTFEE RLAALEGGER ---CVAVGSG
gi|305042|                                                        M TRKQATIAVR SGLND-DEQY GCVVPPIHLS STYNFTGFNE PRA-HDYSRR GNPTRDVVQR ALAELEGGAG ---AVLTNTG
gi|5012317                                                        M TRKQATIAVR SGLND-DEQY GCVVPPIHLS STYNFTGFNQ PRA-HDYSRR GNPTRDVVQR ALAELEGGAG ---AVMTSSG
gi|2889953                                                       MS TRKPATIAVR TGIES-DSQY HAVVPPIYLS TNYGFPAFGE VPK-YDYTRS GNPNRGLLES ALSELESGKG ---AVVTNCG
gi|2122333                                                 MPMSDRHI SQHFETLAIH AGNTA-DPLT GAVVPPIYQV STYKQDGVGG LRGGYEYSRS ANPTRTALEE NLAALEGGRR ---GLAFASG
gi|1567276                                                          MTSIKTKVIH GGIST-DKTT GAVSVPIYQT STYKQNGLGQ PKE-YEYSRS GNPTRHALEE LIADLEGGVQ ---GFAFSSG
gi|1460126                                                          MRGFTTKAVH AGEDPRETLH GDVVHPIHLS VTYWKRSVAE AEEGYVYSRT SNPTRDALEK KLAALENARY ---ALAFSSG
gi|4132487                        MP-KYD- -------NSN ADQWGFETRS IHAGQS--VD AQTSARNLPI YQSTAFVFDS AEHAKQRFAL EDLGPVYSRL TNPTVEALEN RIASLEGGVH ---AVAFSSG
gi|3823323                        MPTKYD- -------NSN ANKWGFETRS IHAGQS--VD SDTSARNLPI YLTSSYVFND AEHAANRFNL SDAGPVYSRL TNPTVAAVEE RLANLEGGVH ---AVLFASG
gi|4140955                     MSSENTDHD- -------TDP SAHWSFETKQ VHAGQH--PD SATNARALPI YQTTSYTFDD TTHAAALFGL EVPGNIYTRI GNPTTDVVEQ RIAALEGGVA ---ALFLSSG
gi|5402556                         MTNDP- -------AEQ AREWSFETKQ IHVGQV--PD TTTNARALPI YQTTSYTFRD TAHAAALFGL AEPGNIYTRI MNPTQDAVEQ RIAALEGGVA ---ALLLASG
gi|1589603                                      MSE ERKFGFETLQ VHAGQV--AD PTTGSRAVPI YQTTSYVFKN ADHAANLFQL KEPGNVYTRI MNPTTDVFEQ RVAALEGGVA ---GLATASG
gi|4886609                          MIEV- -------TTL ENNYHFETEQ IHAGQE-NPD PATGARAVPI YQTSSFVFKD AKEAAGRFAL TEPGNIYGRL TNPTNSAFET RIAALEGGSS ---AISLASG
gi|2302488                             M- -------TNP EKKYAFETLQ LHAGQE-EPD SATGARAVPI YQTTSFVFKD AKQAAGRFAL TDAGNIYGRL TNPTNSAFEA RVAALEGGTA ---AISLASG
gi|4565771                                        M PRNYKPETIA LHGGQS--PD PSTLSRAVPI YQTTSYVFKN TEHAAKLFGL QEFGNIYTRI MNPTTDVLEQ RIAALEGGVA ---ALATASG
gi|4883842  LNEKEG---- -KENRGSQTN ESQTKESRE- -------KKK SESLGVSTLA VHAGAK--PD PATGARSVPI YQTAAYVFKD AEEAADLFGL RKEGNIYTRL MNPTTDVFEK RIAALEGGIG ---ALAVASG
gi|3386538                                        MSQRFETLQ LHAGQS--PD SATNARAVPI YQTSSYVFND AEHGANLFGL KEFGNIYTRL MNPTTDVFEK RVAALEGGVA ---ALATASG
gi|3324025                                       MTSHRFETLQ LHAGQE--PD PSTNSRAVPI YQTSSYVFND AEHGANLFGL KEFGNIYTRL MNPTTDVFEK RVAALEGGVA ---ALATASG
gi|3993764                                          MRNETLA IHAGYE--PD PTTHAVAVPI YQTASYAFDS ADHGAALFNL ETEGYRYSRI ANPTTSVLEK RVAELEGGVG ---ALAVASG
   Consensus .....       .........  .........        ... .......t.. .hag...... ..t.a...pi ..t..y.f.. .......f.. ...g..Y.R. .NPT....E. r.aaLEgg..    ala.asG
```

Fig. 2(to be continued)

```
gi|2852454  MCASTVMLLA LVPAGGHIVT TTDCYRKTRI FMENFLPKLG ITVTVIDPAD IAGLEAAVNE FKVSLFFTES PTNPFLRCVD IELVSKICHK RGTLVCIDGT F-ATPLNQKA LALGADLVVH SATKYIGGHN
gi|8439541  MCASTVMFLA LVPAGGHIVT TTDCYRKTRV FIETILPKMG ITATVIDPAD MGALELTLNQ KKVDLFFTES PTNPFLRCVD IELVSKLCRE KGALVCIDGT F-ATPLNQKA LALGADLVVH SATKFLGGHN
gi|4322948  MCASTVMLLA LVPAGGHIVT TTDCYRKTRI FIETILPKMG ITATVIDPAD VGALELALNQ KKVNLFFTES PTNPFLRCVD IELVSKLCHE KGALVCIDGT F-ATPLNQKA LALGADLVLH SATKFLGGHN
gi|2198853  MYAAAAMLSA LVPAGGHIVT TTDCYRKTRI YMETELPKRG ISMTVIRPAD MDALQNALDN NNVSLFFTET PTNPFLRCID IEHVSNMCHS KGALLCIDST F-ASPINQKA LTLGADLVIH SATKYIAGHN
gi|5379875  MNALTTVMLA LLPSGSHVIL TDDGYRRTRQ FVRTMLARLG VTHSVVAAAD PAAIAAAIEP GRTRLIVTEA PTNPYLRVID LAAVASIARE HRIKTLIDAT F-ATPYNMRP LEYGIDLVVH SCSKYLAGHN
gi|4191271  MAAVAAAVLC QVKAGDHIVA ARALFGSCRW VVETLAPKYG VECTLVDGRD LANWEKAIRP -NTKVFFLES PTNPTLEVVD ISGVAKLANQ VGAKLIVDNV F-ATPLFQKP LELGAHIVVY SATKHIDGQG
gi|2698872  MAAILAVVMS LCSAGDHVLV SQSVFGSTIS LFEKYFKRFG VEVDYVPLVD LTGWEKAIKA -NTKLLIVES PSNPLAELVD ITALSEIAHA QGAMLVVDNC F-STPALQQP LKLGADIVFH SATKFIDGQG
gi|1559830  MSAILALVMS LCSSGDHVLV SRSVFGSTIS LFDKYFKRFG IQVDYPPLSD LAAWEAACKP -NTKLFFVES PSNPLAELVD IAALAEIAHA KGALLAVDNC F-CTPALQQP LKLGADVVIH SATKYIDGQG
gi|5380333  MAAIASTAFG LLKAGDHVVC SRSVFGNTTL LFQNYLAKFG VPTTFVGLTD YDGWAAAIRP -ETRFLFIET PSNPLTEIAD IPRLAEIAHS RGCLLVVDNC F-CTPALQRP LALGADIVIH SATKYLDGQG
gi|305042|  MSAIHLVTTV FLKPGDLLVA PHDCYGGSYR LFDSLAKRGC YRVLFVDQGD EQALRAALAE -KPKLVLVES PSNPLLRVVD IAKICHLARE VGAVSVVDNT F-LSPALQNP LALGADLVLH SCTKYLNGHS
gi|5012317  MSAIMLVCTV FLRPGDLLVA PHDCYGGSYR LFDSLSKRGA FRVKFVDQGD THALNAALAE -KPKLVLVES PSNPLLRVVD IAAICQAARE AGAVSVVDNT F-LSPALQKP LELGADLVVH SCTKYLNGHS
gi|2889953  TSALNLWVSA FLGPDDLIVA PHDCYGGTYR LFNTRANKGD FKVQFVDQSD EQALEAALAK -KPKLILLET PSNPLVRVVD IAAVCEKAKQ VGALVAVDNT F-LTPVYQKP LELGADFVIH STTKYINGHS
gi|2122333  LAAEDCLLRT LLRPGDHVVI PNDAYGGTFR LFAKVATRWG VEWSVADTSD AAAVRAALTP -KTKAVWVET PSNPLLGITD IAQVAQVARD AGARLVVDNT F-ATPYLQQP LALGADVVVH SLTKYMGGHS
gi|1567276  LAGIHAVL-S LFSAGDHIIL ADDVYGGTFR LMDKVLTKTG IIYDLVDLSN LDDLKAAFKE -ETKAIYFET PSNPLLKVLD IKEISAIAKA HDALTLVDNT F-ATPYLQQP IALGADIVLH SATKYLGGHS
gi|1460126  LAAEATVILA LLREGDHIVA FDDLYGGTKR LFQRVMARFG IKVTYVDARD PGNIEKALTP -ATRMIWLET PTNPLLKLTD IRAAAEIASR AGAYLVVDNT F-ATPYFQRP LELGADIVVH SATKYLSGHS
gi|4132487  QAATTNAILN LAGAGDHIVT SPRLYGGTET LFLITLNRLG IDVSFVENPD DPESWQAAVQ PNTKAFFGET FANPQADVLD IPAVAEVAHR NSVPLIIDNT I-ATAALVRP LELGADVVVA SLTKFYTGNG
gi|3823323  MAAETAAILN IARAGSHIVS SPRIYGGTET LFAVTLARLG IETTFVENPD DPASWEAAVQ DNTVALYGET FANPQADVLD IPAIAEVAHK HQVPLIVDNT L-ATAALVRP LELGADVVVA SLTKFYTGNG
gi|4140955  QAAETFAILN LASAGDHIVS SPRLYGGTYN LFHYSLAKLG IEVSFVEDPD DLDSWQAAVR PNTKAFFAET ISNPQIDILD IPGVSGVAHA NGVPLIVDNT I-ATPYLIQP FAHGADIVVH SATKYLGGHG
gi|5402556  QAAETFAILN IAGAGDHIVS SPRLYGGTYN LFHYTLPKLG IEVSFVEDPD DLEQWRAAIR PNTKAFYGET ISNPQNHIFD IPGIAEVAHG EGIPLIVDNT V-ATPYLIRP LEHGADIVVH SATKYLGGHG
gi|1589603  LAAITYAILN VASAGDEIVA ASTLYGGTYE LFGVTLKKLG IKVVFV-DPD NPENIRKAIN DRTKAVYGET IGNPRINVLD IEAVAKIAHE NKIPLIIDNT F-GTPYLIRP IEFGADIVVH SATKFIGGHG
gi|4886609  AAAVTAAILN VAGAGDHIVS ASTLYGGTYE LFSETLKKLG IETTFV-DPD DPANFQAAIK ENTKAIFFES LGNPAINIVD FQAVAAIAKK YKIITVVDST F-ATPFLVRP FDYGIDIVVH SATKFIGGHG
gi|2302488  AAAITAAILN VAGAGDHIVA ASTLYGGTVE LFSETLKKLS IDTTFV-DPD DPENFEKAIQ DNTKVIFFES LGNPKINIID FEAVAAIAKK HGIISIVDST F-ATPFLTRP LEYGIDVVVH SATKFIGGHG
gi|4565771  QAAETLALLN IVEAGQEIVA SSSLYGGTYN LLHYTFPKLG IKVHFV-DPS DPENFRKAVN DKTRAFYAET LGNPKLDTLN LEAIAKVAHD SEVPLIIDNT L-PSPYLVNP IEHGADIVVH SLTKFLGGHG
gi|4883842  MAAITTALLT FTRPGDEIIS GDKLYGGTYE LFNYTFPKLG RTVKFV-DVG KPEEFKNAIS EKTKAIYVES IGNPGLDIPD FEKLAEIAHG AGIPFVVDNT V--SPLILRP IDHGVDIVVY SATKFIGGHG
gi|3386538  QSAQFLAITN CMQAGDNFVS TSFLYGGTYN QFKVQFPRLG IDVRFA-DGD DVESFAAQID DKTKGLYVEA MGNPRFNIPD FEGLSALAKE RGIPLIVDNT LGACGALMRP IDHGADVVVE SATKWIGGHG
gi|3324025  QSAQFLAVSN CMQAGDNFVS TSFLYGGTYN QFKVQFPRLG INVKFA-DGD HIDSFASQID SQTKAIYIES MGNPRFNIPD FSALSDLAKA NKIPLIVDNT LGAAGALIRP LEHGADIVVQ SATKWIGGHG
gi|3993764  QAALHFAFVN LADHGGNIVS VPQLYGTTHT LLSHILPRQG ITGRFA-ASD KPDDIAKLID EGTRAVFCET IGNPAGNVCD IEAIADVAHR AGVPLIVDNT V-ATPILFKP IAYGADVVVH SLTKFLGGHG
 Consensus  .aA.....l. l..aGdhi!. ....%ggt.. lf...l.k.g i.v.fv...# ......a... ..tk....Et p.NP.....D i.a.a..a.. .g..l.!Dnt f atp.lq.p l.lGaD.Vvh SaTKy.gGhg
```

Fig. 2(to be continued)

```
gi|2852454   DVLAGCICGS LK-------- ---------- ---------- ---------- ---------- ---LVS-EIR NLHHVLGGTL NPNAAYLIIR GMKTLHLRVQ QQNSTAFRMA EILEAHPKVS HVYYPGLPSH
gi|8439541   DVLAGCISGP EK-------- ---------- ---------- ---------- ---------- ---LVS-VIR NLHHILGGAL NPNAAYLIIR GMKTLHLRVQ QQNSTALRMA EILEAHPKVK HVYYPGLPSH
gi|4322948   DVLAGCISGP LK-------- ---------- ---------- ---------- ---------- ---LVS-EIR NLHHILGGAL NPNAAYLIIR GMKTLHLRVQ QQNSTALRMA EILEAHPKVR HVYYPGLQSH
gi|2198853   DVIGGCVSGR DE-------- ---------- ---------- ---------- ---------- ---LVS-KVR IYHHVVGGVL NPNAAYLILR GMKTLHLRVQ CQNNTALRMA QFLEEHLKIA RVYYPGLPSH
gi|5379875   DLLAGVIIGR PP-------- ---------- ---------- ---------- ---------- ---IIS-ALR ETQGILGGIC DPHTAYLLGR GLKTFALRME RHNHNGQAVA EFLARHPRVS RVHYPGLPEH
gi|4191271   ACLGGVVLAD KA-------- ---------- ---------- ---------- ---------- ---WIDENLH DYFRHTGPAM SPFNAWTLLK GIETLPLRVR QQTENAAKIA DFLAEQGKVA KVIYPGRKDH
gi|2698872   RCMGGVVAGR TE-------- ---------- ---------- ---------- ---------- ---QMK-EVV GFLRTAGPTL SPFNAWIFTK GLETLRLRMR AHCESAQALA EWLEQQDGVE KVHYAGLPSH
gi|1559830   RGMGGVVAGR GE-------- ---------- ---------- ---------- ---------- ---QMK-EVV GFLRTAGPTL SPFNAWLFLK GLETLRIRMQ AHSASALALA EWLERQPGIE RVYYAGLPSH
gi|5380333   RCVGGAIVGG RE-------- ---------- ---------- ---------- ---------- ---LLDAEIY PFLRTGGPSM SPFNAWVFLK GLETLNLRMK AHCENALGLA RWLEAQPWVE RVHYPGLASH
gi|305042|   DVVAGVVIAK DP-------- ---------- ---------- ---------- ---------- ---DVVTELA WWANNIGVTG GAFDSYLLLR GLRTLVPRME LAQRNAQAIV KYLQTQPLVK KLYHPSLPEN
gi|5012317   DVVAGAVIAK DV-------- ---------- ---------- ---------- ---------- ---DTITELA WWANNIGVTG AAFDSYLLLR GLRTLSPRMA AAQRNAQQVV EFLQTQPLVK ALYHPSLPNN
gi|2889953   DVIGGVVITK TE-------- ---------- ---------- ---------- ---------- ---AHAEELA WWGNCIGATG TPFDSYMTLR GIRTLGARMR VHEESSQHVL NYLQQQALVA KIYHPSLPDH
gi|2122333   DVVGGALIVG DQ-------- ---------- ---------- ---------- ---------- ---ELGEELA FHQNAMGAVA GPFDSWLVLR GTKTLAVRMD RHSENATKVA DMLSRHARVT SVLYPGLPEH
gi|1567276   DVVAGLVTTN SK-------- ---------- ---------- ---------- ---------- ---ELASEIG FLQNSIGAVL GPQDSWLVQR GIKTLALRME AHSANAQKIA EFLETSKAVS KVYYPGLNSH
gi|1460126   DLLAGAVMVN SS-------- ---------- ---------- ---------- ---------- ---EVYEKLK YHQNAVGAVP PPLDSYLLMR SLKTLALRME WHQHNAMKIA EHLESHPKVD RVIYPGLSTH
gi|4132487   SGLGGVLIDG GKFDWTVEKD GKPVFPYFVT PDAAYHGLKY ADLGA----- ---------- -PAFGLKVRV GLLRDTGSTL SAFNAWAAVQ GIDTLSLRLE RHNENAIKVA EFLNNHEKVE KVNFAGLKDS
gi|3823323   SGLGGVLIDG GNFDWTVTRN GEPIFPDFVT PDPAYHGLKY SDLGA----- ---------- -PAFGLKARV GLLRDTGAAP SPLNAWITAQ GLDTLSLRVQ RHNENALAVA QFLANHEKVA KVNYAGLPDS
gi|4140955   SAIAGVIVDG GTFDWTQGR- ----FPGFTT PDPSYHGVVF AELGP----- ---------- -PAYALKARV QLLRDLGSAA APFNAFLVAQ GIETLSLRME RHVANAQRVA EFLASHDGVL SVNYAGLPGS
gi|5402556   AAIAGAIVDG GTFDWTRGR- ----FPGFTE PDPSYHGVKY ADLGA----- ---------- -PAYALKARV QLLRDLGSAV SPFNAFLISQ GLETLSLRIE RHVQNAQAVA EFLAGRSEVT QVFYAGLPSS
gi|1589603   TTIGGIIVDG GKFDWRASGK ----FPDFTT PDKSYNGLIY ADLGA----- ---------- -PAFALKARV QLLRNTGATL SPQSAFYFLQ GLESLSLRVQ KHVDNTRKVV EFLKNHPKVS WINYPELEES
gi|4886609   TTIGGVVVEK GDFDYQASNR ----YPDFVN PTASYNGLIW SDLKG----- ---------- -AAFTTKIRA EHLRDTGATL SPQSAWYFLQ GLETLSLRIE RHVLNTRKIV DYLQQHSKVA WVSYPETENS
gi|2302488   TTLGGVVIEK GDFDYEASGR ----YPDFTT PTPSYNGIVW SDLKG----- ---------- -GAFVTKIRA EHLRDTGATL SPQSAFYLLQ GLETLSLRVE RHVENTRKIV SYLNNNDKVA WVSYPELDDS
gi|4565771   TSIGGIIVDS GKFNWG-NGK ----FKNFTE PDPSYHGLKF WEVFG----- --KFEPFGGV NIAYIIKAKV QGLRDMGASI SPFNAWQILQ GVETLPLRMR KHSENALAVA EYLTKHTKVS WVNYPGLKMD
gi|4883842   TSIGGVIVDS GNFYWK-PEK ----FPEICE PDPGYHGLKY KEAFG----- --K------- -AAFIAKARV QFIRDTGACI SPFNSFLFTL GLETLPLRMK KHCDNALEIA KFLEKHPKVS WVSYPGLESH
gi|3386538   TSLGGVIVDA GTFNWGNGK- ----FPLLSQ PSAAYHGLVH WDAFGFGSDV CKMLGVPDNR NVAFALRARV EGLRDWGPAV SPFNSFLLLQ GLETLSLRVE RHTENAMALA TWLATHPNVE HVSYPGLSSD
gi|3324025   TSLGGVIVDG GTFDWGNGK- ----FPLMSQ PSAAYHGLIH WDAFGFGSDI CKMLGVPDNR NIAFALRARL ECLRDWGPAL SPFNSFLLLQ GLETLSLRIE RHCFNASELA KWLKSHPKIA SVNFPGLPED
gi|3993764   TTLGGAIVDS GRFDWAKQPE R---FPAFNQ PDHSYHGMVY AERFG----- ---------- PTAYVERARS IYQRTMGSVL SPFNAFLLLQ GIETVALRME RHVENARKVA EFLRDDPRVA WVNYTGFPDS
  Consensus  ...gGv.... .......... .   ...... .......... .....        ........ .......... ...r..G... spf.a.l.l. GleTL.lRm. .h..na...a e.L..hp.!. .!.ypglp.h
```

Fig. 2(to be continued)

```
           521                                                                                                                      642
gi|2852454  PEHELAKRQM TG--FGGVVS FEIDGDIETT IKFVDSLKIP YIAPSFGGCE SIVDQPAIMS YWDLPQEERL KYGIKDNLVR FSFGVEDFED VKADILQALE AI
gi|8439541  PEYHLAKKQM TG--FGGVVS FEVDGDLLTT AKFVDALRIP YIAPSFGGCE SIVDQPAIMS YWDLSQSDRA KYGILDNLVR FSFGVEDFED VKADVLQALD SI
gi|4322948  PEHHIAKKQM TG--FGGAVS FEVDGDLLTT AKFVDALKIP YIAPSFGGCE SIVDQPAIMS YWDLSQSDRA KYGIMDNLVR FSFGVEDFDD LKADILQALD SI
gi|2198853  PEHHIAKSQM TG--FGGVVS FEVAGDFDAT RKFIDSVKIP YHAPSFGGCE SIIDQPAIMS YWD-SKEQRD IYGIKDNLIR FSIGVEDFED LKNDLVQALE KI
gi|5379875  PDHAVARAQM RG--FGGVVS FEVVGDLQSA MAVVDRLRLP YIAPSFGGVE SLVEQPALMS YYELSSEERL AVGIRDNLIR LSCGIEDVSD LIADLQQALA DE
gi|4191271  PQAEIIAKQM TG--GSTLVC FELKGGKEAA FALQNALEII KISNNLGDTK SLITHPATTT HKNLTEEARA ELGISPGTVR LSAGIEDTDD LIEDFARGLT KVSA
gi|2698872  PQHELAKRQM SG--FGAVVS FEVKGGKEGA WRFIDATRVI SITTNLGDSK TTIAHPATTS HGRLSPQERE AAGIRDSLIR VAVGLEDVAD LQADLARGLA AL
gi|1559830  PQHELARRQQ SG--FGAVVS FDVKGGRDAA WRFIDATRMV SITTNLGDTK TTIAHPATTS HGRLSPEDRA RAGIGDSLIR VAVGLEDLDD LKADMARGLA AL
gi|5380333  PQHELAARQQ SG--FGGIVS FEVKGGQEAA WRLIDSTRLL SITGNLGDAK TTITHPATTT HGRLSPEARA AAGIADGLIR IAVGLENLAD IQADLARFA
gi|305042|  QGHEIAARQQ KG--FGAMLS FELDGDEQTL RRFLGGLSLF TLAESLGGVE SLISHAATMT HAGMAPEARA AAGISETLLR ISTGIEDGED LIADLENGFR AANKG
gi|5012317  PGHDIARRQQ SG--FGAMLS FELDGDEDTL RRFLASLELF TLAESLGGVE SLISHAATMT HAGMAPEARA AAGISETLLR ISTGIEDGDD LVADLDRAFQ AAAKR
gi|2889953  PGHEIAKKQQ SG--FGSMLS FEFAGSFEQL KVFVKELELF SLAESLGGVE SLICHPASMT HRAMGEEALA EACVSHQLLR LSVGLEDAQD LIADLDQAFV KAAQ
gi|2122333  PGHEVAAKQM KA--FGGMVS FRVEGGEQAA VEVCNRAKVF TLGESLGGVE SLIEHPGRMT HASAAGSALE ---VPADLVR LSVGIENADD LLADLQQALG
gi|1567276  PGHEIAKKQM SA--FGGMIS FELTD-ENAV KDFVENLSYF TLAESLGGVE SLIEVPAVMT HASIPKELRE EIGIKDGLIR LSVGVEAIED LLTDIKEALE KK
gi|1460126  PQHSLAKRQM TG--YSGMLS FELKGGAGAA VRFVESLRII ALAESLGGVE SLVEIPSLMT HASIPREERL KKGITDGLVR LSVGIEDLED LIEDIEQALK KT
gi|4132487  PWYATKEKLG LKY-TGSVLT FEIKGGKDEA WAFIDALKLH SNLANIGDVR SLVVHPATTT HSQSDEAGLA RAGVTQSTVR LSVGIETIDD IIADLEGGFA AI
gi|3823323  PWYPVKEKLG FDY-TGSVLS FDVKGGKNEA WRFIDALKLH SNLANVGDVR SLVVHPATTT HSQSEESALL AAGINQATIR LSVGIESIDD IIADLTAGFD AI
gi|4140955  PWHERAKKLA PKG-TGAVLS FELAGGVEAG KAFVNALKLH SHVANIGDVR SLVIHPASTT HAQLSPAEQL STGVSPGLVR LAVGIEGIDD ILADLELGFA AAAKFGANAD TSGADPHAVA AF
gi|5402556  PWYERGRSLA PKG-TGAVIG FELAGGVDAG KRFVEALTLH SHVANIGDVR SLVIHPASTT HSQLTPEEQL NAGVTPGLVR LAVGIEHIDD ILADLEAGFA AANG
gi|1589603  PYKELANKYL PKG-AGSIFT FGIKGGLEAG KRFINSVKLF SLLANVADAK SLVIHPSSTT HAELNEEEQK AAGVTPDMIR LSIGVEDAED LIWDLNQALE QA
gi|4886609  KYKALAKKYF PKG-SGSIFT FGLKSGESGA ETLIDNLKIF SLLANVGDAK SLIIHPKSTT HAQLNDQQLK AVGITLDLIR VSIGIENVED LIADLDQALT KV
gi|2302488  PYKPLADKYF KNG-VGSIFT FGLKAGEAGA EKLINNLDIF SLLANVGDAK SLIIHPKSTT HAQLNDEQLQ AAGITPDLIR ISIGIENVDD LINALDEALS YVK
gi|4565771  KNYSLAKKYH KKDLYGAILG FGIKGGAVEA KKFIDGLELF SLLANVGDAK SLVIHPASTT HQQLTPEEQL SAGVTPDFVR LSVGLENIED ILFDLEEALK KV
gi|4883842  CSYELAKKYL KSG-YGALIG FGIKGGAREC KKFIEGLEIF SHLANIGDAK SLVIHPASTT HEQLSKEEQL ACGVTEDFIR LSIGIEDAKD LISDIKQALS EV
gi|3386538  PYHAAAKKYL TGRGMGCMLM FSLKGGYDDA VRFINSLQLA SHLANVGDAK TLVIHPASTT HQQLSEDEQA SAGVTPTMVR VSVGLEHIDD IKADFEQALA
gi|3324025  PYHKRAKQYL SGRGMGCMLM FSLKGGFDDA VSFINALKLS SHLANVGDSK TLVIHPASTT HQQLSEDEQK SAGVSPTMVR VSVGLEHIDD IKSDFEQALD QIT
gi|3993764  PYYPLVQKYL DGR-ASSLFT FGIKGGMEAG KAFYDSLKLI TRLVNIGDAK SLACHPASTT HRQMSAEQQR QAGVLPETIR LSIGIEHIAD IIEDLDQALA QAC--GSQPR LAAAE
  Consensus p.he.a.kq. .g...g...s Fe.kgg...a ..f...l.l. s...n.Gd.. sl!.hPa.tt h..l..e... .aG!...l.R lsvGiEd..D l.aDl.qal. .....
```