



(86) Date de dépôt PCT/PCT Filing Date: 2010/07/30
(87) Date publication PCT/PCT Publication Date: 2011/02/03
(45) Date de délivrance/Issue Date: 2017/11/28
(85) Entrée phase nationale/National Entry: 2012/01/12
(86) N° demande PCT/PCT Application No.: EP 2010/061094
(87) N° publication PCT/PCT Publication No.: 2011/012693
(30) Priorités/Priorities: 2009/07/30 (US61/230,076);
2009/07/30 (EP09166815.2)

(51) Cl.Int./Int.Cl. *C12N 9/00* (2006.01),
C12N 15/00 (2006.01), *C12N 9/88* (2006.01),
C12P 7/02 (2006.01), *C12P 7/56* (2006.01)

(72) Inventeurs/Inventors:
VOELKER, FRANCOIS, FR;
DUMON-SEIGNOVERT, LAURENCE, FR;
SOUCAILLE, PHILIPPE, FR

(73) Propriétaire/Owner:
METABOLIC EXPLORER, FR

(74) Agent: SMART & BIGGAR

(54) Titre : METHYLGLYOXAL SYNTHETASE (MGS) MUTANTE POUR LA PRODUCTION D'UN AGENT BIOCHIMIQUE
PAR FERMENTATION
(54) Title: MUTANT METHYLGLYOXAL SYNTHASE (MGS) FOR THE PRODUCTION OF A BIOCHEMICAL BY
FERMENTATION

(57) **Abrégé/Abstract:**

The present invention concerns a method for the production of a biochemical selected among lactic acid, acetol and 1,2-propanediol, comprising culturing a microorganism modified for an improved production of the biochemical selected among lactic acid, acetol and 1,2-propanediol in an appropriate culture medium and recovery of the desired biochemical which may be further purified wherein the microorganism expresses a methylglyoxal synthase (MGS) enzyme which activity is not inhibited by orthophosphate. The present invention concerns a mutant methylglyoxal synthase (MGS) comprising at least one amino acid residue in the protein sequence of the parent enzyme replaced by a different amino acid residue at the same position wherein - the mutant enzyme has retained more than 50% of the methylglyoxal synthase activity of the parent enzyme and - the methylglyoxal synthase activity of the mutant MGS is not inhibited by orthophosphate as compared to the parent enzyme.



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

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
3 February 2011 (03.02.2011)

PCT

(10) International Publication Number
WO 2011/012693 A1

(51) International Patent Classification:

C12N 9/00 (2006.01) *C12P 7/02* (2006.01)
C12N 9/88 (2006.01) *C12P 7/56* (2006.01)
C12N 15/00 (2006.01)

(21) International Application Number:

PCT/EP2010/061094

(22) International Filing Date:

30 July 2010 (30.07.2010)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

09166815.2 30 July 2009 (30.07.2009) EP
 61/230,076 30 July 2009 (30.07.2009) 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*): **VOELKER, François** [FR/FR]; 16 rue Yves Montand, F-42210 Montrond Les Bains (FR). **DUMON-SEIGNOVERT, Laurence** [FR/FR]; 12 rue Marguerite de Valois, F-63430 Pont Du Chateau (FR). **SOUCAILLE, Philippe** [FR/FR]; Chant du Coucou, F-31450 Deyme (FR).

(74) **Agent**: **TETAZ, Franck**; Cabinet Regimbeau, 139 rue Vendôme, F-69477 Lyon Cedex 06 (FR).

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

(84) **Designated States** (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:— *of inventorship (Rule 4.17(iv))***Published:**— *with international search report (Art. 21(3))*— *with sequence listing part of description (Rule 5.2(a))*

(54) **Title**: MUTANT METHYLGLYOXAL SYNTHASE (MGS) FOR THE PRODUCTION OF A BIOCHEMICAL BY FERMENTATION

(57) **Abstract**: The present invention concerns a method for the production of a biochemical selected among lactic acid, acetol and 1,2-propanediol, comprising culturing a microorganism modified for an improved production of the biochemical selected among lactic acid, acetol and 1,2-propanediol in an appropriate culture medium and recovery of the desired biochemical which may be further purified wherein the microorganism expresses a methylglyoxal synthase (MGS) enzyme which activity is not inhibited by orthophosphate. The present invention concerns a mutant methylglyoxal synthase (MGS) comprising at least one amino acid residue in the protein sequence of the parent enzyme replaced by a different amino acid residue at the same position wherein - the mutant enzyme has retained more than 50% of the methylglyoxal synthase activity of the parent enzyme and - the methylglyoxal synthase activity of the mutant MGS is not inhibited by orthophosphate as compared to the parent enzyme.



WO 2011/012693 A1

Mutant methylglyoxal synthase (MGS) for the production of a biochemical by fermentation**TECHNICAL FIELD**

The present disclosure concerns a method for the production of a biochemical selected among lactic acid, acetol and 1,2-propanediol, comprising culturing a microorganism modified for an improved production of the biochemical selected among lactic acid, acetol and 1,2-propanediol in an appropriate culture medium and recovery of the desired biochemical which may be further purified wherein the microorganism expresses a methylglyoxal synthase (MGS) enzyme which activity is not inhibited by orthophosphate.

The present disclosure also relates to a mutant methylglyoxal synthase (MGS) comprising at least one amino acid residue in the protein sequence of the parent enzyme replaced by a different amino acid residue at the same position wherein

- the mutant enzyme has retained more than 50% of the methylglyoxal synthase activity of the parent enzyme and
- the methylglyoxal synthase activity of the mutant MGS is not inhibited by orthophosphate as compared to the parent enzyme.

BACKGROUND

Methylglyoxal synthase (MGS) was discovered and identified as the first enzyme of the methylglyoxal bypass in *E. coli*. MGS was later found in a wide range of organisms including Gram-negative as well as Gram-positive bacteria and yeast (Cooper (1984)). Methylglyoxal bypass can serve as an alternative pathway for the conversion of triosephosphates to pyruvate during the catabolism of glucose (Cooper and Anderson (1970), Cooper (1984)). The Embden-Meyerhoff-Parnas (EMP) pathway or glycolysis involves the conversion of the triosephosphate glyceraldehyde-3-phosphate (G3P) to pyruvate whereas the methylglyoxal bypass starts from the second triosephosphate, dihydroxyacetone phosphate (DHAP), that is converted to pyruvate via the intermediates methylglyoxal (MG) and lactate.

MGS, which was first purified and characterized in *E. coli* (Hopper and Cooper (1971 and 1972)) catalyses the conversion of 1 mole of DHAP to 1 mole of MG and 1 mole of orthophosphate (Pi). MGS is very specific for DHAP, which seems to be the only substrate accepted by the enzyme with a good affinity (Affinity constant K_m varied from 0.2 to 0.47 mM). Several inhibitors of the enzyme were identified: phosphoenolpyruvate (PEP), 3-phosphoglycerate, Pi and pyrophosphate (PPi).

The recent identification of the gene coding for MGS in *E. coli* (*yccG*, then renamed *mgsA*) allowed easier production and characterization of recombinant MGS after

cloning and overexpression of the *mgsA* gene (Tötemeyer et al (1998)). A refined characterization of the enzyme was proposed (Saadat and Harrison (1998)) and the inhibition by the most potent inhibitor, Pi, was further investigated: Pi acted as an allosteric inhibitor of the enzyme, meaning that in the presence of phosphate, a higher amount of DHAP was necessary for the enzymatic reaction to proceed (See also characterization of native MGS from *E. coli* given in Example 2). Several MGS mutants (on positions D20, D71, D91, D10 and H98), always impairing the catalytic rate of the enzyme were characterized and a catalytic mechanism was proposed (Saadat and Harrison (1998), Marks et al (2004)). The three dimensional structure of MGS from *E. coli* was determined after crystallisation of the enzyme (Saadat and Harrison (1999 and 2000)). MGS is a homohexamer with 6 identical units of 17 kDa. Phosphate can bind to the active site of MGS and a hypothesis for the transmission of allosteric information through the salt bridges between the monomers was proposed, although no clear evidence was given.

Production of several products of interest, lactate, acetol and 1,2-propanediol, can result from the catabolism of different carbon substrates (glucose, fructose, sucrose, glycerol) through the methylglyoxal bypass and especially through MGS.

The routes for catabolism of methylglyoxal have been investigated in bacteria (Ferguson et al, 1998) to understand the detoxification of this compounds but also for purposes of production of 1,2-propanediol. Three pathways that can lead to the production of lactate from methylglyoxal have been identified in *E. coli* :

- The first one is the glutathione dependent glyoxalase I-II system (encoded by *gloA* and *gloB* genes) which converts methylglyoxal into D-lactate in two steps (Cooper, 1984).
- The second is the glutathione independent glyoxalase III enzyme which catalyses the conversion of methylglyoxal into D-lactate in one step (Misra et al, 1995).
- The third system is the degradation of methylglyoxal by methylglyoxal reductases, resulting either in acetol or in D- or L-lactaldehyde (Cameron et al, 1998, Bennett and San, 2001). L-lactaldehyde can be further converted to L-lactate by the action of aldehyde dehydrogenases e.g. by the enzymes encoded by the *aldA* or *aldB* genes (Grabar et al, 2006).

Lactate produced by one of the three systems can be further transformed into pyruvate by D- or L-lactate dehydrogenases. These enzymes, in contrast to fermentative lactate dehydrogenases, are flavin-linked membrane-bound proteins that are activated only under aerobic conditions (Garvie, 1980). D- and L-lactate dehydrogenases are coded respectively by the *dld* and *lldD* (or *lctD*) genes in *E. coli* (Rule et al, 1985, Dong et al, 1993).

Acetol or lactaldehyde produced by the third system can be converted to 1,2-propanediol by several enzymatic activities, especially glycerol dehydrogenase (encoded

by *gldA* gene) or 1,2-propanediol oxidoreductase (encoded by *fucO* gene) in *E. coli* (Altaras and Cameron, 2000).

1,2-propanediol or propylene glycol, a C3 dialcohol, is a widely-used chemical. It is a component of unsaturated polyester resins, liquid detergents, coolants, anti-freeze and de-icing fluids for aircraft. Propylene glycol has been increasingly used since 1993-1994 as a replacement for ethylene derivatives, which are recognised as being more toxic than propylene derivatives.

1,2-propanediol is currently produced by chemical means using a propylene oxide hydration process that consumes large amounts of water. Propylene oxide can be produced by either of two processes, one using epichlorhydrin, and the other hydroperoxide. Both routes use highly toxic substances. In addition, the hydroperoxide route generates by-products such as *tert*-butanol and 1-phenyl ethanol. For the production of propylene to be profitable, a use must be found for these by-products. The chemical route generally produces racemic 1,2-propanediol, whereas each of the two stereoisomers (*R*)-1,2-propanediol and (*S*)-1,2-propanediol are of interest for certain applications (*e.g.* chiral starting materials for specialty chemicals and pharmaceutical products).

Acetol or hydroxyacetone (1-hydroxy- 2-propanone) is a C3 keto alcohol. This product is used in vat dyeing process in the textile industry as a reducing agent. It can advantageously replace traditional sulphur containing reducing agents in order to reduce the sulphur content in wastewater, harmful for the environment. Acetol is also a starting material for the chemical industry, used for example to make polyols or heterocyclic molecules. It possesses also interesting chelating and solvent properties.

Acetol is currently produced mainly by catalytic oxidation or dehydration of 1,2-propanediol. New processes starting from renewable feedstocks like glycerol are now proposed (see DE4128692 and WO 2005/095536). Currently, the production cost of acetol by chemical processes reduces its industrial applications and markets.

The disadvantages of the chemical processes for the production of 1,2-propanediol and acetol make biological synthesis an attractive alternative. MGS is the mandatory first step from central metabolism for the production of these two compounds. Processes for the production of 1,2-propanediol or acetol using different microorganism, *Clostridium sphenoides* (DE3336051), *Klebsiella pneumoniae* (WO 2004/087936), recombinant yeast (WO 99/28481) or recombinant *E. coli* (WO 98/37204) have been disclosed. Alternative approaches for the production of 1,2-propanediol or acetol have also been disclosed (WO 2005/073364, WO 2008/116852, WO 2008/116848, WO 2008/116849, WO 2008/116851)).

Lactic acid or lactate and its derivatives have a wide range of applications in the food, pharmaceutical, leather and textile industries. Recently, polylactic acid (PLA) has been developed as a renewable, biodegradable and environmentally friendly plastic and

therefore, the demand for lactate is expected to expand. Lactate can be produced either by a chemical synthesis or by a biological process. However, only a biological process is able to produce the desired stereoisomer, D- or L-lactate with high optical purity, which is an important characteristic for many of its end uses. Physical properties and biodegradation rate of PLA can be controlled by manipulating the ratio of the chiral substrates, D- and L-lactate. Therefore, availability of biological processes for the production of optically pure D- and L-lactate is a prerequisite for high quality polymer synthesis.

Lactic acid bacteria are natural producers of lactate and some can be found to be specific for the D- or L-form. These bacteria have been traditionally used for the production of lactate as specialty chemical (*e.g.* in US 2004/0005677). However, with the emergence of lactate as commodity chemical for PLA synthesis, more efficient and cost-effective processes are needed. Alternative biocatalysts able to growth in mineral salt medium and to use a range of different sugar substrates are investigated. Yeasts and *E. coli* combine these characteristics with the availability of a wide range of genetic tools for metabolic engineering. Use of these catalysts for the production of lactic acid has been described in WO 03102201, WO 03102152 and US 2005/0112737 for yeast strains and in EP 1760156 and WO 2005/033324 for *E. coli* strains. These production processes for D- or L-lactate in microorganisms rely on the reduction of pyruvate produced by the catabolism of sugars by NADH-dependent lactate dehydrogenases, generally under anaerobic conditions. The methylglyoxal bypass with the three pathways for the degradation of MG mentioned above can serve as an alternative non-fermentative pathway for the production of lactate, as described in PCT/EP2009/053093.

According to the allosteric inhibition of MGS by Pi, the conditions necessary for the enzyme to be active would be a high concentration of its substrate DHAP or a low concentration of Pi. When Pi is limiting in the environment, G3P dehydrogenase cannot continue to work without one of its substrate and therefore G3P and hence DHAP will accumulate, filling the two conditions for efficient work of MGS. Under these conditions, methylglyoxal bypass will replace glycolysis for catabolism of triosephosphates. When Pi is abundant, glycolysis will operate because the concentration of Pi would be too high and the concentration of DHAP too low for MGS to be active (Cooper (1984), Fergusson et al (1998)). This mechanism allows the microorganism to cope with different situations with regards to Pi. However, concerning the production of metabolites in the methylglyoxal bypass when these molecules are the end-products of the metabolism, the two parallel pathways, glycolysis and methylglyoxal bypass will have to work together: glycolysis to ensure the supply of precursors and energy for growth and methylglyoxal bypass for the synthesis of the wanted products. In this case, a MGS enzyme that has lost its inhibition by phosphate would be a clear advantage.

The inventors have identified new mutant MGS that had lost allosteric inhibition by phosphate, while keeping most of their specific activity for the conversion of DHAP into MG, as demonstrated in Example 2 by the characterization of purified enzymes. Use of these mutants is a key element in the design of more efficient processes for the production of the products of the methylglyoxal bypass, particularly acetol, 1,2-propanediol and lactate.

SUMMARY

The present disclosure concerns a method for the production of a biochemical selected among lactic acid, acetol and 1,2-propanediol, comprising culturing a microorganism modified for an improved production of the biochemical selected among lactic acid, acetol and 1,2-propanediol in an appropriate culture medium and recovery of the desired biochemical which may be further purified wherein the microorganism expresses a methylglyoxal synthase (MGS) enzyme which activity is not inhibited by orthophosphate.

The present disclosure concerns a mutant methylglyoxal synthase (MGS) enzyme, comprising an amino acid sequence selected from the group consisting of SEQ ID NO 1, SEQ ID NO 2 and SEQ ID NO 3.

The present disclosure concerns a mutant methylglyoxal synthase (MGS) comprising at least one amino acid residue in the protein sequence of the parent enzyme replaced by a different amino acid residue at the same position wherein

- the mutant enzyme has retained more than 50% of the methylglyoxal synthase activity of the parent enzyme and
- the methylglyoxal synthase activity of the mutant MGS is not inhibited by orthophosphate as compared to the parent enzyme.

The disclosure also concerns a DNA sequence comprising a sequence coding for the mutant MGS of the disclosure and a microorganism expressing such MGS which activity is not inhibited by orthophosphate, particularly a microorganism comprising a gene coding for the mutant MGS of the invention.

DETAILED DESCRIPTION

In the present application, terms are employed with their usual meaning, except when precised otherwise.

Microorganisms

A “microorganism” means all kind of unicellular organisms, including procaryotic organisms like bacteria, and eucaryotic organisms like yeasts. Preferentially, the microorganism is selected among the group consisting of bacteria, yeasts and fungi, more preferentially selected among *Enterobacteriaceae*, *Bacillaceae*, *Streptomyetaceae*, *Clostridiaceae* and *Corynebacteriaceae*. More preferentially, the microorganism is a species of *Escherichia*,

Klebsiella, *Pantoea*, *Salmonella*, *Bacillus*, *Streptomyces*, *Clostridium* or *Corynebacterium*. Even more preferentially, the microorganism is selected among the group consisting of *Escherichia coli*, *Klebsiella pneumoniae*, *Thermoanaerobacterium thermosaccharolyticum*, *Clostridium sphenoides* or *Saccharomyces cerevisiae*.

5 As used herein, the term “modified microorganism” or “modified” or “recombinant” refer to a host cell that has a modification of its genome, e.g., as by addition of nucleic acid not naturally occurring in the organism or by a modification of nucleic acid naturally occurring in the host cell.

10 A “microorganism modified for an improved production of the biochemical selected among lactic acid, acetol and 1,2-propanediol” is a microorganism in which pathways to favour the production of the desired biochemical by conversion of a simple source of carbon have been modified. The microorganism modified for such improved production produces more of the desired biochemical than a native, unmodified microorganism.

15 The preferred biosynthetic pathways for the production of lactic acid, acetol and propanediol with the microorganism of the disclosure are represented on Figure 2. The person skilled in the art shall identify the enzymatic activities related to the pathway to be promoted and the other enzymatic activities to be attenuated.

20 Microorganisms modified for the improved production of lactic acid, acetol and propanediol by conversion of methylglyoxal are also disclosed in Cameron et al, 1998, Bennett and San, 2001, Ko et al, 2005 and WO 99/28481, WO 98/37204, WO 2005/073364, WO 2008/116852, WO 2008/116848, WO 2008/116849, WO 2008/116851, PCT/EP2009/053093, which content is incorporated herein by reference.

In the case of yeasts, the following modifications of the host organism are preferred:

- 25 ▪ attenuation of expression of at least one of the following genes : TPI1, NDE1, NDE2, GUT2, GPD1, GPD2, PDC1, PDC2, PDC5, PDC6, GLO1
- enhancement of expression of GRE3 gene.

In the microorganisms of the invention, the DNA sequence coding for a mutant MGS of the disclosure may be introduced in a vector for the expression and translation of the mutant MGS. It can also be integrated in the chromosome of the said microorganism.

30 Integration of the DNA sequence can be done either entirely, or simply by introducing in the native gene of the microorganism, the mutation in the coding sequence by replacing the nucleotide(s) coding for the amino acid to be changed by the nucleotide(s) coding for the amino acid of the mutated protein.

35 Total, partial or specific nucleotides replacement in a gene of a microorganism is well known in the art of genetic engineering, including Sambrook J et al., Molecular cloning : a

laboratory manual, Cold Spring Harbour Press, New York (2001), Ausubel FM et al., Current protocols in molecular biology, John Wiley and sons, New York (1999), Adams A et al., Methods in yeast genetics, Cold Spring Harbour Press, New York (1997).

The microorganism of the disclosure may additionally comprise a gene coding for an YqhD enzyme which catalytic efficiency towards NADPH is increased.

An YqhD enzyme “*which catalytic efficiency toward NADPH is increased*” means that the catalytic efficiency towards NADPH of the YqhD enzyme expressed in the microorganism is higher than the catalytic efficiency towards NADPH of the native YqhD enzyme of the same microorganism. The catalytic efficiency is defined as the ratio between the catalytic constant (Kcat) and the Michaelis constant (Km). Increase of catalytic efficiency of YqhD enzyme means that the Kcat of the enzyme is increased or that the Km of the enzyme is decreased. In a preferred embodiment the Kcat of the YqhD enzyme is increased and the Km of the YqhD enzyme is decreased.

Preferably, the catalytic efficiency towards NADPH of the YqhD enzyme is higher than the efficiency of the native YqhD enzyme of *E. coli*.

Such enzyme has preferably an enzymatic activity of at least 50% of the activity of the YqhD of *E. coli*, more preferably at least 60% of the activity of the YqhD of *E. coli*.

Particularly, the YqhD enzyme is a mutant YqhDS enzyme wherein

- the mutant enzyme has retained more than 50% of the YqhD activity of the parent enzyme and
- the catalytic efficiency toward NADPH of the mutant YqhD is increased as compared with the catalytic efficiency toward NADPH of the parent enzyme.

Preferably the mutant YqhD comprises at least a mutation selected among the group consisting of G149E, G149S and A286T, and combinations thereof. The aminoacids positions are given by reference to the YqhD sequence of *E. coli*. The person skilled in the art shall find the corresponding aminoacids in sequences from other organisms by standard techniques of sequence alignment.

The microorganism of the disclosure may additionally comprise a gene coding for a glycerol dehydrogenase (GlyDH) enzyme with reduced inhibition of its activity by NAD⁺ and/or its substrate and/or its product.

“*The inhibition of which activity by NAD⁺ and/or its substrate and/or its product is reduced*” means that the inhibition of the activity of the GlyDH enzyme expressed in the microorganism is less inhibited than the activity of the native GlyDH enzyme of the same microorganism. The inhibition of the activity of the GlyDH enzyme can be defined by the Inhibition Concentration 50 (IC50) or the Inhibition Constant (Ki) or any other techniques known by the skilled person. The reduced inhibition of the activity of the GlyDH enzyme means that the

IC50 or the Ki of the GlyDH enzyme of the disclosure is higher than the IC50 or the Ki of the native GlyDH enzyme. The skilled person knows the relation between IC50 and Ki and their meaning on the activity of enzyme among the classic Michaelis-Menten kinetics.

Preferably, the activity of the GlyDH enzyme is less inhibited than the native GlyDH enzyme of *E. coli*. In a preferred embodiment, the enzyme activity is less inhibited for at least two members of the group consisting of NAD⁺, the enzyme's substrate and the enzyme's product. More preferably the enzyme activity is less inhibited by the three of NAD⁺, its substrate and its product.

The enzyme "*substrate*" is dihydroxyacetone, hydroxyacetone, methylglyoxal, lactaldehyde, glyceraldehyde, glycolaldehyde and derivatives thereof.

The enzyme "*product*" is the molecule obtained from the selected substrate by reduction of the carbonyl function.

For the production of 1,2-propanediol, the substrate is hydroxyacetone and the product is 1,2-propanediol.

Particularly the GlyDH enzyme is a mutant enzyme, wherein

- the mutant enzyme has retained more than 50% of the activity of the parent enzyme and
- the glycerol dehydrogenase activity of the mutant GlyDH is less inhibited by NAD⁺ and/or by its substrate as compared to the parent enzyme and/or by its product as compared to the parent enzyme.

Preferably, the mutant GlyDH preferably comprises at least a mutation selected among the group consisting of A160T and T120N, and combinations thereof. The aminoacids positions are given by reference to the GldA sequence of *E. coli*. The person skilled in the art shall find the corresponding aminoacids in sequences from other organisms by standard techniques of sequence alignment.

Methylglyoxal synthase (MGS) enzyme

The disclosure concerns a methylglyoxal synthase (MGS) which activity is not inhibited by orthophosphate a microorganism comprising the same and a method for the production of a desired biochemical by fermentation of said microorganism on a culture medium comprising a simple source of carbon.

"*Not inhibited by orthophosphate*" or "*lacking inhibition by orthophosphate*" means that no inhibition by orthophosphate is identified in an activity assay, when activity of the enzyme is studied in presence of orthophosphate. Such an activity assay is well known in the art and can be carried out as disclosed in Example 2.

In addition, kinetics of the MGS enzyme of the disclosure follows Michaelis-Menten kinetics regardless of the presence or absence of orthophosphate. Kinetics of the native enzyme

follow a Michaelis-Menten model only in the absence of orthophosphate. The presence of orthophosphate makes the kinetic profile (specific activity over substrate concentration) of the native enzyme to become sigmoidal, which denotes the allosteric inhibition by orthophosphate.

Such enzyme has preferably a methylglyoxal synthase activity of at least 50% of the activity of the methylglyoxal synthase of *E. coli*.

The enzymes may be obtained by various ways known to the person skilled in the art.

A first approach consists in screening native enzymes of various organisms for their lack of inhibition by orthophosphate.

A second approach consists in inducing mutation(s) in enzymes of known organisms and selecting the enzymes for their lack of inhibition by orthophosphate. Mutations may be induced by methods known in the art such as subjecting the microorganism to mutagenic agents. Another method to induce mutations is to grow the microorganism under selection pressure, with high levels of orthophosphate and identify the microorganism growing under such conditions and select the enzymes obtained for their lack of inhibition by orthophosphate.

Other methods are also known in the art to obtain mutations by shuffling DNA from various origins and select the proteins encoded by the shuffled DNA so obtained based on their methylglyoxal synthase activity and their lack of inhibition by orthophosphate.

In a particular embodiment of the disclosure, the inventors obtained several mutants MGS retaining their methylglyoxal synthase activity and lacking inhibition by orthophosphate by selecting strains modified for an improved production lactic acid, acetol and/or 1,2-propanediol cultured under selection pressure as disclosed in WO 2005/073364 and as shown in Example 1.

The disclosure concerns particularly mutant methylglyoxal synthase (MGS) comprising at least one amino acid residue in the protein sequence of the parent enzyme replaced by a different amino acid residue at the same position wherein

- the mutant enzyme has retained more than 50% of the activity of the parent enzyme and
- the methylglyoxal synthase activity of the mutant MGS is not inhibited by orthophosphate as compared to the parent enzyme.

“*Mutant*” means that a mutation was introduced in a protein sequence by a person. According to the disclosure, a mutant enzyme is an enzyme comprising at least one amino acid difference with a parent enzyme. In the mutant enzyme of the invention, any change in amino acids may be introduced, either by directed mutagenesis or random mutagenesis, but also using chimerical enzymes comprising parts of a second enzyme replacing corresponding parts of the parent enzyme.

The “*parent enzyme*” is the enzyme prior mutation. The parent enzyme may be of any origin, natural, isolated from another organism or synthetic.

The method for determining that the mutated MGS has retained “*more than 50%*” is well known in the art and disclosed in Example 2.

Indeed, the skilled person shall choose the level of desired activity according to the final use of mutant MGS. Indeed, when a high activity is necessary, the skilled person will choose a mutant having more than 80% of activity, compared to the non mutated parent enzyme, more preferably more than 90% of activity. In other cases, selecting a mutant MGS with an activity around and above 50% compared to the parent enzyme may prevent additional modifications in a microorganism, like modifying the promoter to lower the level of expression of the enzyme.

The inventors found that in a medium comprising up to 3 mM orthophosphate, the enzymatic activity of the mutant enzyme is not inhibited by orthophosphate when the corresponding parent enzyme is inhibited. The amount of orthophosphate was chosen provided that the enzymatic reaction generally occurs in a microorganism.

According to the invention, “*not inhibited by orthophosphate as compared to the parent MGS*” is understood to be at orthophosphate concentrations compatible with the concentrations of orthophosphate in a microorganism growing on a medium comprising orthophosphate at concentrations compatible with the growth of the microorganism.

In a preferred embodiment, the mutant MGS of the disclosure comprises at least one amino acid residue of an identified region in the native parent MGS replaced by a different amino acid residue at the same position.

The inventors have identified mutants wherein at least one amino acid residue of one of the following Conserved Regions (CR) in the native parent MGS has been replaced by a different amino acid residue at the same position, the three Conserved Regions CR1, CR2 and CR3 being identified below:

-Xa1-Leu-Xa2-Xa3-His-Asp-Xa4-Xa5-Lys- (CR1)

wherein

Xa1 represents Ala and Val, preferably Ala,

Xa2 represents Val and Ile,

Xa3 represents Ala and Ser, preferably Ala,

Xa4 represents Ala, Arg, Asn, Gln, Glu, His, Lys, Met and Ser, preferably His and Lys, and

Xa5 represents Arg, Cys, Gln, Lys, Met and Tyr, preferably Cys and Lys

-Asp-Xa6-Xa7-Xa8-Xa9-X10-X11-His-X12-X13-Asp-X14- (CR2)

wherein

Xa6 represents Asp and Pro, preferably Pro,

Xa7 represents Leu and Met, preferably Leu,

Xa8 represents Asn, Glu, Ser and Thr, preferably Asn and Thr,

Xa9 represents Ala, Asn, Pro, Ser and Val, preferably Ala,

X10 represents Ala, Leu, Gln, Lys, Met and Val, preferably Gln and Val,

X11 represents Ala and Pro, preferably Pro,

5 X12 represents Asp and Glu,

X13 represents Ala, Pro and Val, preferably Pro, and

X14 represents Ile and Val, preferably Val

-X15-X16-X17-X18-Pro-X19-X20-X21-X22- (CR3)

wherein

10 X15 represents Ile, Leu and Val, preferably Val,

X16 represents Arg, Gln, His, Trp and Tyr, preferably Trp and Tyr,

X17 represents Ala, Asn, Arg, Asp, Gln, Glu, Gly, Lys and Ser, preferably Asn,

X18 represents Ile, Leu and Val, preferably Ile,

X19 represents Cys, His, Ile, Leu, Met and Val, preferably Leu and Val,

15 X20 represents Ala and Val, preferably Ala,

X21 represents Cys, Ile, Leu, Met and Thr, preferably Thr, and

X22 represents Asn and Thr, preferably Asn.

These conserved regions can be identified in different MGS enzymes by simple sequence alignment using standard sequence alignment tools such as ClustalW2, Kalign, 20 MAFFT, MUSCLE or T-coffee, all available on the website <http://www.ebi.ac.uk/>. A sequence alignment of several MGS of different species is given in Figure 1.

Amino acids numbers in the present application are given by reference to the proteins of *E. coli*.

It can be found in figure 1 that CR1 correspond to amino acids 15 to 23 of *E. coli* 25 MGS, CR2 correspond to amino acids 91 to 102 of *E. coli* MGS and CR3 correspond to amino acids 111 to 119 of *E. coli* MGS.

According to the invention, the mutant MGS can have at least one mutation in one of CR1, CR2 or CR3. It can have at least two mutations in CR1 and CR2, in CR1 and CR3 or in CR2 and CR3. It can also have at least three mutations in CR1, CR2 and CR3.

30 “At least” in such context means that the mutated enzyme may have other mutations, but not related to the identified Conserved Regions CR1, CR2 and CR3. These other non identified mutations have no substantial impact on the mutated enzyme of the invention, provided that:

- the mutant enzyme has retained more than 50% of the methylglyoxal synthase activity 35 of the parent enzyme and
- the methylglyoxal synthase activity of the mutant MGS is not inhibited by orthophosphate as compared to the parent enzyme.

In preferred embodiments, the amino acid residue in the Conserved Regions CR1 to CR3 in the parent MGS replaced by a different amino acid residue at the same position in the mutant MGS is selected among the group consisting of amino acid Xa4 in CR1, amino acid Xa9 in CR2 and amino acid X19 in CR3 and combinations thereof (CR1 & CR2, CR1 & CR3, CR2 & CR3 and CR1 & CR2 & CR3).

Xa4 correspond to amino acid 21 in the MGS sequence of *E. coli*. Xa9 corresponds to amino acid 95 in the MGS sequence of *E. coli*. X19 corresponds to amino acid 116 in the sequence of *E. coli*.

Particularly, the mutated MGS of the disclosure comprises at least one of the mutations selected among the group consisting in H21Q, A95V, V116L, and combinations thereof, the aminoacid positions being given by reference to the MGS sequence of *E. coli*.

More preferably, the mutated MGS of the disclosure comprises at least one of the following amino acid sequence in conserved regions CR1 to CR3

- CR1: Ala Leu Val Ala His Asp **Gln** Cys Lys
- CR2: Asp Pro Leu Asn **Val** Val Pro His Asp Pro Asp Val
- CR3: Val Trp Asn Ile Pro **Leu** Ala Thr Asn

the amino acid residue marked in bold and underlined corresponding to the amino acid in the mutant MGS different from the amino acid in the parent MGS.

Particularly, the mutant MGS of the disclosure has at least 50% sequence identity compared to the MGS sequence of *E. coli*, provided that it comprises at least one of the following mutations in CR1 and/or CR2:

- CR1: Ala Leu Val Ala His Asp **Gln** Cys Lys
- CR2: Asp Pro Leu Asn **Val** Val Pro His Asp Pro Asp Val
- CR3: Val Trp Asn Ile Pro **Leu** Ala Thr Asn.

Sequence identity is defined after sequence alignment of the MGS sequence of *E. coli* with the protein sequence to be compared using CLUSTALW2 available on the EBI website (see above) with default parameters. The sequence identity is then calculated with the sequence alignment by the ratio of the number of identical amino acids at the same position with the total number of amino acids in the reference sequence (*E. coli*).

Preferably, the mutant MGS has at least 70% sequence identity.

In most preferred embodiments, the mutant MGS of the disclosure comprises the sequence selected among the group consisting of MGS identified in SEQ ID NO 1, SEQ ID NO 2 and SEQ ID NO 3.

DNA, vectors, genes

The present disclosure also concerns a DNA sequence comprising a sequence coding for the mutant MGS of the invention. The sequence coding for the mutant MGS of the disclosure is not a limiting factor by itself. The skilled person can easily obtain the sequence of a native MGS from a microorganism and introduce in the coding sequence the mutation(s) to be introduced in the protein by changing one or more appropriate nucleotide.

The skilled person can also perform a mutagenesis in the sequence of a microorganism, and isolate the mutated DNA sequence by standard methods.

Mutations can be introduced by site-directed mutagenesis by usual methods like Polymerase Chain Reaction (PCR, see Sambrook J et al., Molecular cloning : a laboratory manual, Cold Spring Harbour Press, New York (2001), Ausubel FM et al., Current protocols in molecular biology, John Wiley and sons, New York (1999), Adams A et al., Methods in yeast genetics, Cold Spring Harbour Press, New York (1997)), or by random mutagenesis techniques, such as use of mutagenic agents (Ultra-Violet rays or chemical agents like nitrosoguanidine (NTG) or ethylmethanesulfonate (EMS)) or use of PCR techniques (DNA shuffling or error-prone PCR).

The person skilled in the art can also prepare synthetic genes with preferred codons selected for an improved expression in a specific organism. Codons usages by various organisms are well known in the art and several companies are proposing the manufacture of synthetic genes with codon optimization.

The sequence of the disclosure can be isolated, consisting in the coding sequence as defined above, or within a gene comprising regulatory elements upstream and downstream the coding sequence for its expression in a specific organism.

The sequence can also be present in a vector, for its replication (replication vector) or for the expression and translation of the mutated protein of the disclosure in a microorganism (expression vector). Such vectors are known in the art and not a limiting factor for the definition of the invention.

Said genes and vectors are also part of the invention.

Preferably, the DNA sequence of the disclosure is in a microorganism with regulatory elements allowing expression and translation of the mutated MGS of the invention.

Production of lactic acid, acetol, or 1,2-propanediol

The disclosure also concerns a method for the production of a biochemical selected among lactic acid, acetol and 1,2-propanediol by fermentation comprising culturing a microorganism of the invention, modified for an improved production of lactic acid, acetol and/or 1,2-propanediol and recovery of the biochemical.

In a particular embodiment, the recovered lactic acid and/or acetol and/or 1,2-propanediol is purified.

Methods for the purification of lactic acid, acetol and 1,2-propanediol are known in the art and described in Datta and Henry, 2006, Wasewar, 2005, US5076896 and WO 2007/074066.

Advantageously, the production is done by fermentation in a batch, fed-batch or continuous process, according to processes known to the person skilled in the art of microorganisms fermentation. Preferably, the production is done by fermentation in a fed-batch process.

Culture medium and Carbon source

In the production method of the disclosure, the microorganism is cultured on an appropriate culture medium.

An “*appropriate culture medium*” means a medium of known molecular composition adapted to the growth of the micro-organism. In particular, said medium contains at least a source of phosphorus and a source of nitrogen. Said appropriate medium is for example a mineral culture medium of known set composition adapted to the bacteria used, containing at least one carbon source. Said appropriate medium may also designate any liquid comprising a source of nitrogen and/or a source of phosphorus, said liquid being added and/or mixed to the source of sucrose. In particular, the mineral growth medium for Enterobacteriaceae can thus be of identical or similar composition to M9 medium (Anderson, 1946), M63 medium (Miller, 1992) or a medium such as defined by Schaefer *et al.* (1999).

The carbon source ‘glucose’ can be replaced in this medium by any other carbon source, in particular by sucrose or any sucrose-containing carbon source such as sugarcane juice or sugar beet juice.

A “*carbon source*” or “*carbon substrate*” means any carbon source capable of being metabolized by a microorganism wherein the substrate contains at least one carbon atom.

Preferably, the carbon source is selected among the group consisting of glucose, sucrose, mono- or oligosaccharides, starch or its derivatives or glycerol and mixtures thereof.

Indeed the microorganisms used in the method of the present disclosure can be modified to be able to grow on specific carbon sources when the non modified microorganism cannot grow on the same source of carbon, or grow at too low rates. These modifications may be necessary when the source of carbon is a byproduct of biomass degradation such as by-products of sugarcane including; filter cake from clarification of raw juice and different kind of molasses.

Figure 2 represents the metabolic pathways for the production of lactic acid, acetol and 1,2-propanediol in the microorganisms of the invention.

Figure 3 represents the analysis on SDS 4-15% gradient polyacrylamide gel the different purification steps of the protein mgsA* V116L. Lane 1: Molecular weight marker, Lane 2: Crude extract, Lane 3, supernatant of the 1st ammonium sulphate precipitation, Lane 4, pellet of the 2nd ammonium sulphate precipitation, Lane 5, Resource Q pool, Lane 6, MonoQ pool, Lane 7, superdex 200 pool, Lane 8, final pool with the BSA.

EXAMPLES

10

Example 1: Evolution of 3 modified strains of *E. coli* MG1655 in chemostat culture and identification of 3 mutant MGS enzymes in the evolved clones:

The construction of the strains *E. coli* MG1655 *lpd** $\Delta tpiA$, $\Delta pflAB$, $\Delta adhE$, *ldhA::Km*, $\Delta gloA$, $\Delta aldA$, $\Delta aldB$, Δedd (strain 1), *E. coli* MG1655 *lpd** $\Delta tpiA$, $\Delta pflAB$, $\Delta adhE$, $\Delta ldhA::Cm$, $\Delta gloA$, $\Delta aldA$, $\Delta aldB$, Δedd (strain 2) and *E. coli* MG1655 *lpd** $\Delta tpiA$, $\Delta pflAB$, $\Delta adhE$, $\Delta ldhA$, $\Delta gloA$, $\Delta aldA$, $\Delta aldB$, Δedd , $\Delta arcA$, $\Delta ndh::Km$ (strain 3) were previously described in patent application WO 2005/073364 for strain 1 and in patent application WO 2008/116852 for strains 2 and 3.

To evolve them toward improved 1,2 propanediol production, the 3 strains were cultivated in continuous culture, either under anaerobic conditions, or under microaerobic conditions (1% oxygen) in the culture medium MPG (given in patent application WO 2008/116852) with 0.42 or 0.84 g/l sodium nitrate, with excess glucose (from 20 g/l initially with addition if the glucose becomes exhausted). The temperature was set at 37°C, the pH was regulated at 6.5 by addition of base and the dilution rate of the chemostat was set between 0.04 h⁻¹ and 0.08 h⁻¹. The evolution of the strain in the chemostats was followed by the increase of the biomass concentration coupled with the increase of the concentrations of the product, 1,2-propanediol and the co-product acetate, over several weeks. This denoted the improvement of the performances of the strains. When the cultures reached a steady state with no further increase of the concentrations under these conditions, the evolution was done.

The characteristics of the strains before and after evolution were assessed. Single colonies representing individual clones were isolated on Petri dishes. These clones were assessed using the initial strain as control in an Erlenmeyer flask assay, using the same medium MPG used in the chemostat culture, but buffered with MOPS. Among these clones, several presented better 1,2-propanediol specific production rates as compared to the control. The results obtained on the best clone for each condition of evolution are reported in Table 1 to 3 below.

Table 1: Comparison of the best evolved clone obtained after 81 days of evolution under anaerobic conditions with the initial strain

Strain <i>E. coli</i> MG1655 lpd* $\Delta tpiA$ $\Delta pflAB$ $\Delta adhE$ $\Delta ldhA::Km$ $\Delta gloA$ Δald, $\Delta aldB$ Δedd (Strain 1)	Initial strain before evolution (performances measured after 2 days of culture)	Best evolved clone (performances measured after 2 days of culture)
Glucose specific consumption rate (g glucose /g biomass /h)	0.13	0.18 (+ 38%)
1,2-propanediol specific production rate (g 1,2-propanediol /g biomass /h)	0.02	0.06 (+200%)
1,2-propanediol + hydroxyacetone specific production rate (g 1,2-propanediol + hydroxyacetone /g biomass /h)	0.04	0.06 (+ 50%)

5 Table 2: Comparison of the best evolved clone obtained after 66 days of evolution under anaerobic conditions with the initial strain

Strain <i>E. coli</i> MG1655 lpd* $\Delta tpiA$ $\Delta pflAB$ $\Delta adhE$ $\Delta ldhA::Cm$ $\Delta gloA$ Δald, $\Delta aldB$ Δedd (Strain 2)	Initial strain before evolution (performances measured after 2 days of culture)	Best evolved clone (performances measured after 2 days of culture)
Glucose specific consumption rate (g glucose /g biomass /h)	0.12	0.21 (+ 75%)
1,2-propanediol specific production rate (g 1,2-propanediol /g biomass /h)	0.02	0.07 (+250%)
1,2-propanediol + hydroxyacetone specific production rate (g 1,2-propanediol + hydroxyacetone /g biomass /h)	0.04	0.08 (+ 100%)

Table 3: Comparison of the best evolved clone obtained after 132 days of evolution under microaerobic conditions with the initial strain

Strain <i>E. coli</i> MG1655 lpd* $\Delta tpiA$ $\Delta pflAB$ $\Delta adhE$ $\Delta ldhA::Cm$ $\Delta gloA$ Δald, $\Delta aldB$, Δedd, $\Delta arcA$ Δndh (Strain 3)	Initial strain before evolution (performances measured after 2 days of culture)	Best evolved clone (performances measured after 2 days of culture)
Glucose specific consumption rate (g glucose /g biomass /h)	0.15	0.28 (+ 87%)
1,2-propanediol specific production rate (g 1,2-propanediol /g biomass /h)	0.00	0.10
1,2-propanediol + hydroxyacetone specific production rate (g 1,2-propanediol + hydroxyacetone /g biomass /h)	0.04	0.10 (+ 150%)

Specific genes involved in the terminal 1,2-propanediol biosynthetic pathway were sequenced in the 3 best evolved clones of strain 1, strain 2 and strain 3. For each clone, one mutated *mgsA* gene was identified resulting in expression of mutated MGS protein: MgsA*(A95V) (SEQ ID NO 1) for evolved clone of strain 1, MgsA*(H21Q) (SEQ ID NO 2) for evolved clone of strain 2 and MgsA*(V116L) (SEQ ID NO 3) for evolved clone of strain 3.

Example 2 : Production, purification and characterization of native MGS and 3 mutant MGS (H21Q, A95V & V116L)

10

1. Construction of the strains for production of MGS proteins

1.1. Construction of the plasmid for overexpression of *mgsA* : pETTOPO-*mgsA*

15 The plasmid was built to obtain the overexpression of the native protein (without His-tag). The gene *mgsA* (sequence 1025780-1026238) was PCR amplified from genomic DNA of *E. coli* MG1655 using the following oligonucleotides :

- pETTOPO *mgsA* F (consisting of 24 pb) :

caccatggaactgacgactcgca, (SEQ ID NO 4)

20 with - a region (underlined letters) homologous to the sequence (1026238-1026220) of the gene *mgsA*.

- a region (bold) for directional cloning of the fragment in the plasmid pET101.

and

- pETTOPO-N *mgsA* R. (consisting of 23 pb)

25 Ttacttcagacgggtccgcgagat (SEQ ID NO 5)

with a region (underlined letters) homologous to the sequence (1025780-1025802) of the gene *mgsA*.

The fragment amplified was directly cloned into the pET101 from the “Champion pET Directional TOPO Expression Kits” (invitrogen®). The plasmid built was named

30 pETTOPO-*mgsA*.

1.2. Construction of the plasmids for the overexpression of *mgsA**

The three mutant MGS bear the mutations H21Q, A95V or V116L. The plasmids for the overexpression of the three mutant proteins were built by directed mutagenesis in the plasmid pETTOPO-*mgsA* using Quickchange site-directed mutagenesis kit from stratagene®.with the oligonucleotides given in Table 4.

35

Table 4: Oligonucleotides used for the site-directed mutagenesis of *mgsA*

mutant	Names of oligos	Sequence of the oligonucleotide	Modification to create mutation	Homologous to the sequence
A95V	mgsA*A95VmutDir5'	ttctgggatccactaaatgTcgtgccgca cgatcctgacgtCaaagcc (SEQ ID NO 6)	Red capital letter : T instead of C to create A95V mutation Green capital letter: C instead of G to create <i>AatII</i> restriction site without change in protein sequence	1025974 to 1025927
	mgsA*A95VmutDir3'	ggctttGacgtcaggatcgtgcggcag Acatttagtggatcccagaa (SEQ ID NO 7)	Red capital letter : A instead of G to create A95V mutation Green capital letter: C instead of G to create <i>AatII</i> restriction site without disturbing without change in protein sequence	
V116L	mgsA*V116LmutDir5'	gcgacggtatggaacattccgCtcgcG accaacgtggcaacg (SEQ ID NO 8)	Red capital letter : C instead of G to create V116L mutation Green capital letter: G instead of C to create <i>NruI</i> restriction site without change in protein sequence	1025914 to 1025873
	mgsA*V116LmutDir3'	cgttgccacgttggtCgcgaGcggaatg ttccataccgtgc (SEQ ID NO 9)	Red capital letter : G instead of C to create V116L mutation Green capital letter: C instead of G to create <i>NruI</i> restriction site without change in protein sequence	
H21Q	mgsA* H21Q mut dir F	ctggtggcacacgatcaAtgcaaacaG atgctgatgagctgggtg (SEQ ID NO 10)	Red capital letter : A instead of C to create V116L mutation Green capital letter: G instead of A to create <i>AlwNI</i> restriction site without change in protein sequence	1026193 to 1026149
	mgsA* H21Q mut dir R	caccagctcatcagcatCgtttgcaTt gatcgtgtgccaccag (SEQ ID NO 11)	Red capital letter : T instead of G to create V116L mutation Green capital letter: C instead of T to create <i>AlwNI</i> restriction site without change in protein sequence	

The 3 plasmids obtained were named pETTOPO-*mgsA**(A95V), pETTOPO-*mgsA**(V116L) and pETTOPO-*mgsA**(H21Q).

5 1.3. Construction of BL21 star (DE3) $\Delta mgsA::Cm$

To avoid the mix between the mutant proteins expressed by the plasmid and the wild-type one expressed by chromosome, the strain used to carry out the overexpression was deleted for the *mgsA* gene.

1.3.1. Construction of the strain MG1655 $\Delta mgsA::Cm$

10 The gene *mgsA* was inactivated in strain *E. coli* MG1655 by inserting a chloramphenicol antibiotic resistance cassette and deleting most of the gene concerned according to in Protocol 1.

Protocol 1 : Introduction of a PCR product for recombination and selection of the recombinants (FRT system).

The oligonucleotides chosen and given in Table 5 for replacement of a gene or an intergenic region were used to amplify either the chloramphenicol resistance cassette from the plasmid pKD3 or the kanamycin resistance cassette from the plasmid pKD4 (Datsenko, K.A. & Wanner, B.L. (2000)). The PCR product obtained was then introduced by electroporation into the recipient strain bearing the plasmid pKD46 in which the system λ Red ($\gamma, \beta, \text{exo}$) expressed greatly favours homologous recombination. The antibiotic-resistant transformants were then selected and the insertion of the resistance cassette was checked by PCR analysis with the appropriate oligonucleotides given in Table 6.

If they are other modifications in the strain, they were checked with the oligonucleotides given in Table 6.

The resulting strain was named *E. coli* MG1655 $\Delta mgsA::\text{Cm}$

Table 5: Oligonucleotides used for replacement of a chromosomal region by recombination with a PCR product

Region name	Names of oligos	Homology with chromosomal region
mgsA	DmgsA F DmgsA R	1026273 to 1026193 1025837 to 1025758
<i>edd-eda</i>	DedaR DeddF	1930144 to 1930223 1932501 to 1932582
aldA	DaldAR DaldAF	1487615 to 1487695 1486256 to 1486336
aldB	DaldBR DaldBF	3754534 to 3754455 3752996 to 3753075
arcA	DarcAF DarcAR	4637868 to 4637791 4637167 to 4637245
Ndh	DndhF DndhR	1165071 to 1165149 1166607 to 1166528
Ptrc01-gapA	Ptrc-gapAF Ptrc-gapAR	1860800 to 1860762 1860478 to 1860536
gldA	DgldA F DgldA R	4137058 to 4136979 4135955 to 4136034
gloA	DgloA F DgloA R	1725861 to 1725940 1726268 to 1726189
yqhD	DyqhDF DyqhDR	3153377 to 3153456 3154540 to 3154460
Dld	Ddld F Ddld R	2220207 to 2220288 2221919 to 2221838
lldD	DlldP F DlldP R	3777860 to 3777941 3779040 to 3778961

Table 6: Oligonucleotides used for checking the insertion of resistance cassette or the loss of resistance cassette

Region name	Names of oligos	Homology with chromosomal region
mgsA gene	helD F	1025242 to 1025260
	yccT R	1026499 to 1026480
<i>edd-eda</i> genes	edaR	1929754 to 1929777
	eddF	1932996 to 1932968
<i>aldA</i> gene	aldAF	1485877 to 1485901
	aldAR	1487714 to 1487689
<i>aldB</i> gene	aldBF	3752056 to 3752095
	aldBR	3754644 to 3754674
<i>arcA</i> gene	arcAF	4638292 to 4638273
	arcAR	4636854 to 4636874
<i>ndh</i> gene	ndhF	1164722 to 164742
	ndhR	1167197 to 1167177
Ptrc01-gapA	yeaAF	1860259 to 1860287
	gapAR	1861068 to 1861040
<i>gldA</i> gene	talC F	4137144 to 4137121
	yijF R	4135136 to 4135159
gloA gene	gloAF2	1725641 to 1725660
	gloAR2	1726450 to 1726431
yqhD gene	yqhDF	3153068 to 3153100
	yqhDR	3154825 to 3154797
dld gene	dld F	2219708 to 2219729
	dld R	2222343 to 2222323
lldD gene	lldP F	3777394 to 3777414
	lldP R	3779406 to 3779384
mgsA::Km	helDF	1025242 to 1025260
	mgsA R3	1026734 to 1026715

1.3.2. Construction of the strain BL21 star (DE3) Δ mgsA::Cm

5 The deletion of the gene *mgsA* by replacement of the gene by a chloramphenicol resistance cassette in the strain *E. coli* BL21 star (DE3) was performed by the technique of transduction with phage P1.

Protocol 2 : Transduction with phage P1 for deletion of a gene

10 The deletion of the chosen gene by replacement of the gene by a resistance cassette (kanamycin or chloramphenicol) in the recipient *E. coli* strain was performed by the technique of transduction with phage P1. The protocol was in two steps, (i) the preparation of the phage lysate on the strain MG1655 with a single gene deleted and (ii) the transduction of the recipient strain by this phage lysate.

Preparation of the phage lysate

- Seeding with 100 µl of an overnight culture of the strain MG1655 with a single gene deleted of 10 ml of LB + Cm 30 µg/ml + glucose 0.2% + CaCl₂ 5 mM.
- Incubation for 30 min at 37°C with shaking.
- 5 - Addition of 100 µl of phage lysate P1 prepared on the wild type strain MG1655 (approx. 1×10^9 phage/ml).
- Shaking at 37°C for 3 hours until all cells were lysed.
- Addition of 200 µl of chloroform, and vortexing.
- Centrifugation for 10 min at 4500 g to eliminate cell debris.
- 10 - Transfer of supernatant in a sterile tube and addition of 200 µl of chloroform.
- Storage of the lysate at 4°C

Transduction

- Centrifugation for 10 min at 1500 g of 5 ml of an overnight culture of the *E. coli* recipient strain in LB medium.
- 15 - Suspension of the cell pellet in 2.5 ml of MgSO₄ 10 mM, CaCl₂ 5 mM.
- Control tubes: 100 µl cells
100 µl phages P1 of the strain MG1655 with a single gene deleted.
- Tube test: 100 µl of cells + 100 µl phages P1 of strain MG1655 with a single gene deleted.
- 20 - Incubation for 30 min at 30°C without shaking.
- Addition of 100 µl sodium citrate 1 M in each tube, and vortexing.
- Addition of 1 ml of LB.
- Incubation for 1 hour at 37°C with shaking
- Plating on dishes LB + Cm 30 µg/ml after centrifugation of tubes for 3 min at
- 25 7000 rpm.
- Incubation at 37°C overnight.

The antibiotic-resistant transformants were then selected and the insertion of the deletion was checked by a PCR analysis with the appropriate oligonucleotides given in Table 6.

- 30 The resulting strain was named *E. coli* BL21 star (DE3) $\Delta mgsA::Cm$.

1.4. Introduction of plasmids in the strain BL21 star (DE3) $\Delta mgsA::Cm$

- The plasmids pETTOPO-*mgsA*, pETTOPO-*mgsA**(A95V), pETTOPO-*mgsA**(V116L), pETTOPO-*mgsA**(H21Q) were transformed by electroporation in the
- 35 strain *E. coli* BL21 star (DE3) $\Delta mgsA::Cm$ and the strains obtained were named, respectively:

- BL21 star (DE3) $\Delta mgsA::Cm$ pETTOPO-*mgsA*
- BL21 star (DE3) $\Delta mgsA::Cm$ pETTOPO-*mgsA**(A95V)

- BL21 star (DE3) $\Delta mgsA::Cm$ pETTOPO-*mgsA**(V116L)
- BL21 star (DE3) $\Delta mgsA::Cm$ pETTOPO-*mgsA**(H21Q)

2. Production of MGS proteins

5 The four strains BL21 star (DE3) $\Delta mgsA::Cm$ pETTOPO-*mgsA*, BL21 star (DE3) $\Delta mgsA::Cm$ pETTOPO-*mgsA**(A95V), BL21 star (DE3) $\Delta mgsA::Cm$ pETTOPO-*mgsA**(V116L) and BL21 star (DE3) $\Delta mgsA::Cm$ pETTOPO-*mgsA**(H21Q) were cultivated at 37°C under aerobic conditions in 2 l baffled Erlenmeyer flasks with 500 ml LB medium with 2.5 g/l glucose. The flasks were agitated at 200 rpm on an orbital shaker.

10 When the optical density measured at 550 nm reached 0.5 units, the flasks were incubated at 25°C. When the optical density reached 1.2 units, the production of MGS proteins was induced by adding 500 μ M IPTG in the cultures. The biomass was harvested by centrifugation when the cultures reached an optical density above 3.5 units. The supernatant was discarded and the pellet was stored at – 20°C before use.

15

3. Activity assay for MGS

A coupled activity assay, adapted from Hopper and Cooper (1972) was used to determine the enzyme activity. Dihydroxyacetone phosphate (DHAP) is converted to methylglyoxal (MG) by MGS. The formation of MG is coupled to the formation of S-D-lactoylglutathione by the non enzymatic formation to the thio hemi-acetal with glutathione and the subsequent isomerisation of the complex by glyoxalase I. The rate of increase in absorbance at 240 nm, corresponding to S-D-lactoylglutathione formation and thus to MG formation, was measured at 30°C on a spectrophotometer. The standard assay mixture consisted of 1.5mM DHAP, 1.5 mM glutathione, 50 mM Imidazole (pH 7.0), 2 units of yeast glyoxalase I and 30 μ l of MGS sample in a total volume of 1000 μ l. Control assay, lacking the MGS sample was run in parallel and the value measured for the control was subtracted to the assay to take into account non-specific formation of MG or S-D-lactoylglutathione. Initial velocities were measured by following the increase in absorbance att 240 nm over time after addition of MGS sample. A unit MGS activity was defined as the formation of 1 μ mol of MG/min under the conditions of the assay. Specific enzyme activity was expressed as units per mg of protein.

30

4. Purification of the MGS enzymes

The four proteins MgsA, MgsA*(V116L), MgsA*(H21Q), MgsA*(A95V) were purified using the same protocol. This protocol was adapted from Hooper and Cooper (1972).

35

All chromatographic columns were run at room temperature. Fractions were stored at 4°C between purification steps.

4.1. Step 1: Preparation of cell-free extracts

Between 350-400 mg of *E. coli* biomass was resuspended in 70 ml of 50 mM Imidazole, 1mM Potassium Phosphate pH 7, and a protease inhibitor cocktail. Cells were sonicated on ice (Branson sonifier, 70W) in a Rosett cell RZ3 during six cycles of 30 sec with 30 sec intervals. Cells debris were removed by centrifugation at 12000g for 30 min at 4°C. Supernatant was kept as the crude extract.

4.2. Step 2: Ammonium sulphate precipitation

Solid ammonium sulphate (209 g/l) was added to the crude extract on ice. After 15 min of incubation at 4°C, the precipitate was removed by centrifugation at 12000g for 15 min at 4°C and discarded. More ammonium sulphate (111 g/l) was added to the supernatant solution at 0°C. After 15 min of incubation at 4°C, the mix was centrifuged at 12000g for 15 min at 4°C. The supernatant was discarded and the precipitate dissolved in 200 ml of 50 mM Imidazole, 1mM Potassium Phosphate pH 7.

4.3. Step 3: Anionic chromatography pH 7

Using an Akta Purifier (GE Healthcare), the half of ammonium sulphate pellet resuspended in 50 mM Imidazole, 1mM Potassium Phosphate pH 7 (100ml) was loaded onto a 6ml Resource Q column (GE Healthcare) equilibrated with the same buffer. Two runs were realised. For each run, the column was washed with 10 column volumes of the same buffer. Proteins were eluted with a gradient of 20 column volumes from 0 M to 0.5 M sodium chloride. After elution, the column was washed with 1 column volume of a gradient from 0.5M to 1M sodium chloride and with 5 column volumes of 1M sodium chloride. The flow rate of the column was 2 ml/min and 5 ml fractions were collected. The MGS protein was eluted with 150mM sodium chloride. The fractions which contain the MGS protein were pooled and dialysed against 50 mM Imidazole, 1mM Potassium Phosphate, 100mM NaCl pH 8 overnight.

4.4. Step 4: Anionic chromatography pH 8

The dialysed pool was applied to a 1.7ml Mono Q column (GE Healthcare) equilibrated with 50 mM Imidazole, 1mM Potassium Phosphate, 100mM NaCl pH 8. To avoid column overload, 4 runs were done. For each run, the column was washed with 10 column volumes of 50 mM Imidazole, 1mM Potassium Phosphate, 100mM NaCl pH 8. Proteins were eluted with a gradient of 20 column volumes from 0.1 M to 0.5 M sodium chloride. After elution, the column was washed with 1 column volume of a gradient from 0.5M to 1M sodium chloride and with 5 column volumes of 1M sodium chloride. The flow rate of the column was 1.5 ml/min and 2 ml fractions were collected. The MGS protein was eluted with about 200 mM sodium chloride. The fractions which contain the MGS protein were pooled and concentrated to be loaded on a gel filtration column.

4.5. Step 5: Gel filtration

The concentrated fractions from the Mono Q column were loaded onto a Superdex 200 10/300 GL column (GE Healthcare) equilibrated with 50 mM Imidazole, 1mM Potassium Phosphate, 350 mM NaCl pH 7. Four runs were realised. The flow rate of the column was 0.5 ml/min and 0.5 ml fractions were collected. The MGS protein was eluted with about 13,5 ml of buffer. The expression and purification of the mutant MGS were remarkably similar to that of the wild type enzyme. There were no differences in the oligomerisation state between the native mgsA and the mutated mgsA*.

All proteins were stored at 4°C in presence of 0.1 mg/ml BSA to stabilise the protein.

The pool of each purification step was analysed on a SDS 4-15% gradient polyacrylamide gel (Fig. 3). This gel shows the purity increased along the purification step. After the superdex 200 column, the protein was almost 90% pure. The final pool showed two major bands at about 17 kDa corresponding to the protein MgsA and at about 70 kDa corresponding to the BSA used to stabilize the enzyme.

5. Characterization of the MGS enzymes in the absence of orthophosphate

Kinetic constants (K_m , k_{cat} and k_{cat}/K_m) for the four purified enzymes (MgsA, MgsA*(V116L), MgsA*(H21Q) and MgsA*(A95V)) were determined using the activity assay previously described. At least six DHAP concentrations between 0.08 mM and 1.5mM were analysed for each enzyme. For all kinetics, initial velocity was determined in triplicate for all DHAP concentrations.

Before the activity assay, the purified protein stored at 4°C in 50 mM Imidazole, 1mM Potassium Phosphate, 350 mM NaCl, 0.1 mg/ml BSA pH 7 was diluted in 50 mM Imidazole, 10% glycerol, 0.1 mg/ml BSA pH 7.

Kinetic constants of each protein were calculated with the module enzyme kinetics from the software Sigma Plot (Systat Software Inc, San Jose CA). The data sets exhibiting Michaelis-Menten were fitted to a Michaelis-Menten equation. The different kinetic parameters of four MGS were compiled in the Table.

Table 7: kinetics parameters for the MGS enzymes without Pi

	MgsA	MgsA* H21Q	MgsA* V116L	MgsA* A95V
K_m (mM)	$0,19 \pm 0,017$	$0,1746 \pm 0,009$	$0,1606 \pm 0,021$	$0,28 \pm 0,027$
Specific activity ($\mu\text{moles}/\text{min}/\text{mg}$)	1049 ± 24	1056 ± 15	505 ± 16	786 ± 24
k_{cat} (s^{-1})	296 ± 7	298 ± 4	143 ± 4	222 ± 7
k_{cat}/k_m ($M^{-1}s^{-1}$)	$1,54E+06 \pm 1,72E+05$	$1,70E+06 \pm 1,12E+05$	$8,87E+05 \pm 1,44E+05$	$7,84E+05 \pm 9,73E+04$

The kinetics parameters of the four enzymes, the native MGS and the 3 mutant MGS in absence of orthophosphate were very similar. The value for K_m is in good agreement with the value of 0.20 ± 0.03 mM previously reported for the native enzyme (Saadat and Harrison, 1998).

5 Specific activity for each MGS is directly calculated from the k_{cat} value. Specific activity of MgsA and MgsA*(H21Q) were similar. Specific activity of MgsA*(A95V) represented 75% of the specific activity of MgsA. Specific activity of MgsA*(V116L) represented 50% of the specific activity of MgsA. The mutations were not detrimental for the activity of the enzyme.

10

6. Characterization of the MGS enzymes in the presence of orthophosphate

The kinetic parameters (K_m , k_{cat} and k_{cat}/K_m) of the four purified enzymes (MgsA, MgsA*(V116L), MgsA*(H21Q) and MgsA*(A95V)) stored at 4°C in 50 mM Imidazole, 1mM Potassium Phosphate, 350 mM NaCl, 0.1 mg/ml BSA pH 7 were
15 determined in presence of different concentration of potassium phosphate in the activity assay (0.2mM, 0.3mM, 1mM orthophosphate (Pi)). Before the activity assay, the protein is diluted in 50 mM Imidazole, 10% glycerol, 0.1 mg/ml BSA pH 7. To determine precisely these kinetic parameters, all measures of initial velocities were done in triplicate for at least six concentrations in substrate (DHAP) between 0.08 mM and 1.5mM.

20 Kinetic constants of each protein at each concentration of potassium phosphate (0.2mM, 0.3mM, 1mM Pi) were determined with the module enzyme kinetics from the software Sigma Plot. The different kinetic parameters of MGS enzymes were compiled in the Table 8.

Table 8: kinetic parameters for the MGS enzymes in the presence of various concentrations of Pi

MgsA				
	Without Pi	0,2mM Pi	0,3mM Pi	1mM Pi
Km (mM)	0,19 ± 0,017	0,41 ± 0,013	0,57 ± 0,021	> 1,53
Specific activity (μmol/min/mg)	1049 ± 24	1019 ± 24	1104 ± 42	> 592
kcat (s ⁻¹)	296 ± 7	287 ± 7	311 ± 12	> 167
Kcat/Km (M ⁻¹ s ⁻¹)	1,5E+06 ± 1,7E+05	7E+05 ± 3,8E+04	5,4E+05 ± 4,1E+04	> 1,09E+05

MgsA* A95V				
	Without Pi	0,2mM Pi	0,3mM Pi	1mM Pi
Km (mM)	0,28 ± 0,027	0,34 ± 0,077	0,45 ± 0,079	0,51 ± 0,085
Specific activity (μmol/min/mg)	786 ± 24	799 ± 38	901 ± 61	787 ± 52
kcat (s ⁻¹)	222 ± 7	226 ± 11	254 ± 17	222 ± 15
kcat/Km (M ⁻¹ s ⁻¹)	7,8E+05 ± 9,7E+04	6,6E+05 ± 1,8E+05	5,7E+05 ± 1,4E+05	4,31E+05 ± 1,00E+05

MgsA* H21Q				
	Without Pi	0,2mM Pi	0,3mM Pi	1mM Pi
Km (mM)	0,17 ± 0,009	0,22 ± 0,024	0,24 ± 0,016	0,38 ± 0,035
Specific activity (μmol/min/mg)	1056 ± 15	1048 ± 34	1057 ± 21	1056 ± 35
kcat (s ⁻¹)	298 ± 4	295 ± 10	298 ± 6	298 ± 10
Kcat/Km (M ⁻¹ s ⁻¹)	1,7E+06 ± 1,1E+05	1,3E+06 ± 1,8E+05	1,2E+06 ± 1,0E+05	7,9E+05 ± 1,0E+05

MgsA* V116L				
	Without Pi	0,2mM Pi	0,3mM Pi	1mM Pi
Km (mM)	0,16 ± 0,021	0,19 ± 0,026	0,17 ± 0,029	0,47 ± 0,068
Specific activity (μmol/min/mg)	505 ± 16	525 ± 20	502 ± 23	598 ± 33
kcat (s ⁻¹)	143 ± 4	148 ± 6	142 ± 7	169 ± 9
kcat/Km (M ⁻¹ s ⁻¹)	8,9E+05 ± 1,4E+05	7,9E+05 ± 1,4E+05	8,1E+05 ± 1,8E+05	3,6E+05 ± 7,2E+04

5

For the native MgsA, when activity was measured at various concentration of DHAP without phosphate, the enzyme showed standard Michaelis-Menten kinetics. However, the presence of Pi at concentrations of 0.2-0.3 mM caused the response to DHAP to become sigmoidal, and raising the Pi concentration led to an increasingly pronounced sigmoidal response. As a consequence, the Km of the enzyme was sharply increased. This denoted the allosteric inhibition of the MGS enzyme by orthophosphate, as already described in the literature (Saadat and Harrison, 1998).

For the three mutants MgsA*(V116L), MgsA*(H21Q) and MgsA*(A95V), at all concentrations of Pi, the kinetics fitted to a Michaelis-Menten equation and no allosteric inhibition by Pi was found. The Michaelis Menten curves were very similar in absence and in presence of orthophosphate.

To summarize, the properties of the 3 mutant MGS were very similar: the mutant MGS have lost the allosteric inhibition by orthophosphate shown by the native MGS enzyme.

20

Example 3 : Construction of two *E. coli* 1,2-propanediol producer strains expressing wildtype or modified MGS and assessment of 1,2-propanediol production

- 1. Construction of the modified strain *E. coli* MG1655, *mgsA**(H21Q)::Km, Δ *edd-eda*, Δ *gloA*, Δ *aldA*, Δ *aldB*, Δ *arcA*, Δ *ndh* (pME101-VB01-*yqhD**(G149E)-*gldA**(160T)).**

1.1. Construction of a modified strain *E. coli* Δ *gloA*::Cm

The gene *gloA* was inactivated in strain *E. coli* MG1655 by inserting a chloramphenicol antibiotic cassette and deleting most of the gene concerned using the technique described in protocol 1 with the oligonucleotides given in Table 5. The deletion was checked by a PCR analysis with the appropriate oligonucleotides given in Table 6. The resulting strain was named *E. coli* MG1655 Δ *gloA*::Cm.

1.2. Construction of a modified strain *E. coli* Δ *gloA*::Cm Δ *edd-eda*::Km

1.2.1. Construction of a modified strain *E. coli* Δ *edd-eda*::Km

The genes *edd-eda* were inactivated in strain *E. coli* MG1655 by inserting a kanamycin antibiotic cassette and deleting most of the gene concerned using the technique described in Protocol 1 with the oligonucleotides given in Table 5. The deletion was checked by a PCR analysis with the appropriate oligonucleotides given in Table 6. The resulting strain was named *E. coli* MG1655 Δ *edd-eda*::Km.

1.2.2. Construction of a modified strain *E. coli* Δ *gloA*::Cm, Δ *edd-eda*::Km.

The deletion of the genes *edd-eda* by replacement of the genes by a kanamycin resistance cassette in the strain *E. coli* Δ *gloA*::Cm was performed by the technique of transduction with phage P1 according to Protocol 2.

Protocol 2 : Transduction with phage P1 for deletion of a gene

The deletion of the chosen gene by replacement of the gene by a resistance cassette (kanamycin or chloramphenicol) in the recipient *E. coli* strain was performed by the technique of transduction with phage P1. The protocol was in two steps, (i) the preparation of the phage lysate on the strain MG1655 with a single gene deleted and (ii) the transduction of the recipient strain by this phage lysate.

Preparation of the phage lysate

- Seeding with 100 μ l of an overnight culture of the strain MG1655 with a single gene deleted of 10 ml of LB + Cm 30 μ g/ml + glucose 0.2% + CaCl₂ 5 mM.
- Incubation for 30 min at 37°C with shaking.
- Addition of 100 μ l of phage lysate P1 prepared on the wild type strain MG1655 (approx. 1×10^9 phage/ml).
- Shaking at 37°C for 3 hours until all cells were lysed.
- Addition of 200 μ l of chloroform, and vortexing.

- Centrifugation for 10 min at 4500 g to eliminate cell debris.
- Transfer of supernatant in a sterile tube and addition of 200 µl of chloroform.
- Storage of the lysate at 4°C

Transduction

- 5 - Centrifugation for 10 min at 1500 g of 5 ml of an overnight culture of the *E. coli* recipient strain in LB medium.
- Suspension of the cell pellet in 2.5 ml of MgSO₄ 10 mM, CaCl₂ 5 mM.
- Control tubes: 100 µl cells
100 µl phages P1 of the strain MG1655 with a single gene deleted.
- 10 - Tube test: 100 µl of cells + 100 µl phages P1 of strain MG1655 with a single gene deleted.
- Incubation for 30 min at 30°C without shaking.
- Addition of 100 µl sodium citrate 1 M in each tube, and vortexing.
- Addition of 1 ml of LB.
- 15 - Incubation for 1 hour at 37°C with shaking
- Plating on dishes LB + Cm 30 µg/ml after centrifugation of tubes for 3 min at 7000 rpm.
- Incubation at 37°C overnight.

The antibiotic-resistant transformants were then selected and the insertion of the deletion was checked by a PCR analysis with the appropriate oligonucleotides given in Table 6.

The deletion was checked by a PCR analysis with the appropriate oligonucleotides given in Table 6 as well as the other deletions already present in the strain.

The resulting strain was named *E. coli* $\Delta gloA::Cm$, $\Delta edd-eda::Km$.

25

1.3. Construction of a modified strain *E. coli* MG1655 $\Delta gloA$, $\Delta edd-eda$

The antibiotic resistance cassettes were eliminated in the strain *E. coli* $\Delta gloA::Cm$ $\Delta edd-eda::Km$, according to Protocol 3.

Protocol 3 : Elimination of resistance cassettes (FRT system)

30 The chloramphenicol and/or kanamycin resistance cassettes were eliminated according to the following technique. The plasmid pCP20 carrying the FLP recombinase acting at the FRT sites of the chloramphenicol and/or kanamycin resistance cassettes was introduced into the strain by electroporation. After serial culture at 42°C, the loss of the antibiotic resistance cassettes was checked by PCR analysis with the oligonucleotides
35 given in Table 5.

The presence of the modifications previously built in the strain was checked using the oligonucleotides given in Table 6.

The strain obtained was named *E. coli* MG1655 $\Delta gloA$ $\Delta edd-eda$.

1.4. Construction of the modified strain *E. coli* MG1655 Δedd -*eda* $\Delta gloA$, $\Delta aldA::Cm$

1.4.1. Construction of the modified strain *E. coli* MG1655 $\Delta aldA::Cm$

The gene *aldA* was inactivated in strain *E. coli* MG1655 by inserting a
 5 chloramphenicol antibiotic cassette and deleting most of the gene concerned using the
 technique described in protocol 1 with the oligonucleotides given in Table 5. The deletion
 was checked by a PCR analysis with the appropriate oligonucleotides given in Table 6.
 The resulting strain was named *E. coli* MG1655 $\Delta aldA::Cm$.

1.4.2. Construction of the modified strain *E. coli* MG1655 Δedd -*eda* 10 $\Delta gloA$, $\Delta aldA::Cm$

The deletion of the gene *aldA* by replacement of the gene by a chloramphenicol
 resistance cassette in the strain *E. coli* MG1655 Δedd -*eda* $\Delta gloA$ was performed by the
 technique of transduction with phage P1 (Protocol 2).

The deletion $\Delta aldA::Cm$ and the others modifications were checked using the
 15 oligonucleotides described in Table 6.

The resulting strain was named *E. coli* MG1655 Δedd -*eda* $\Delta gloA$, $\Delta aldA::Cm$.

1.5. Construction of the modified strain *E. coli* MG1655 Δedd -*eda*, $\Delta gloA$, $\Delta aldA::Cm$, $\Delta aldB::Km$

1.5.1. Construction of the modified strain *E. coli* MG1655 $\Delta aldB::Km$

The gene *aldB* was inactivated in strain *E. coli* MG1655 by inserting a Kanamycin
 antibiotic cassette and deleting most of the gene concerned using the technique described
 in protocol 1 with the oligonucleotides given in Table 5. The deletion was checked by a
 PCR analysis with the appropriate oligonucleotides given in Table 6.

25 The resulting strain was named *E. coli* MG1655 $\Delta aldB::Km$.

1.5.2. Construction of the modified strain *E. coli* MG1655 Δedd -*eda* $\Delta gloA$, $\Delta aldA::Cm$, $\Delta aldB::Km$

The deletion of the gene *aldA* by replacement of the gene by a Kanamycin
 resistance cassette in the strain *E. coli* MG1655 Δedd -*eda* $\Delta gloA$ $\Delta aldA::Cm$, was
 30 performed by the technique of transduction with phage P1 (Protocol 2).

The deletion $\Delta aldB::Km$ and the others modifications were checked using the
 oligonucleotides described in Table 6.

The resulting strain was named *E. coli* MG1655 Δedd -*eda* $\Delta gloA$, $\Delta aldA::Cm$, $\Delta aldB::Km$.

1.6. Construction of the modified strain *E. coli* MG1655 Δedd - Δeda $\Delta gloA$, $\Delta aldA$, $\Delta aldB$

The antibiotic resistance cassettes were eliminated in the strain *E. coli* MG1655 Δedd - Δeda $\Delta gloA$, $\Delta aldA::Cm$, $\Delta aldB::Km$ according to Protocol 3.

5 The loss of the antibiotic resistance cassettes was checked by PCR analysis with the oligonucleotides given in Table 6. The presence of the modifications previously built in the strain was also checked using the oligonucleotides given in Table 6.

The strain obtained was named *E. coli* MG1655 Δedd - Δeda $\Delta gloA$, $\Delta aldA$, $\Delta aldB$.

10 1.7. Construction of the modified strain *E. coli* MG1655 Δedd - Δeda $\Delta gloA$, $\Delta aldA$, $\Delta aldB$, $\Delta arcA::Km$

1.7.1. Construction of the modified strain *E. coli* MG1655 $\Delta arcA::Km$

The gene *arcA* was inactivated in strain *E. coli* MG1655 by inserting a Kanamycin antibiotic cassette and deleting most of the gene concerned using the technique described
15 in Protocol 1 with the oligonucleotides given in Table 5. The deletion was checked by a PCR analysis with the appropriate oligonucleotides given in Table 6.

The resulting strain was named *E. coli* MG1655 $\Delta arcA::Km$.

1.7.2. Construction of the modified strain *E. coli* MG1655 Δedd - Δeda $\Delta gloA$, $\Delta aldA$, $\Delta aldB$, $\Delta arcA::Km$

20 The deletion of the gene *arcA* by replacement of the gene by a Kanamycin resistance cassette in the strain *E. coli* Δedd - Δeda $\Delta gloA$, $\Delta aldA$, $\Delta aldB$, $\Delta arcA::Km$, was performed by the technique of transduction with phage P1 (Protocol 2)

The deletion $\Delta arcA::Km$ and the others modifications were checked using the oligonucleotides described in Table 6.

25 The strain obtained was named *E. coli* MG1655 Δedd - Δeda $\Delta gloA$, $\Delta aldA$, Δald , $\Delta arcA::Km$.

1.8. Construction of the modified strain *E. coli* MG1655 Δedd - Δeda $\Delta gloA$, $\Delta aldA$, $\Delta aldB$, $\Delta arcA::Km$, $\Delta ndh::Cm$

30 1.8.1. Construction of the modified strain *E. coli* MG1655 $\Delta ndh::Cm$

The gene *ndh* was inactivated in strain *E. coli* MG1655 by inserting a chloramphenicol resistance cassette and deleting most of the gene concerned using the technique described in Protocol 1 with the oligonucleotides given in Table 5. The deletion was checked by a PCR analysis with the appropriate oligonucleotides given in Table 6.

35 The resulting strain was named *E. coli* MG1655 $\Delta ndh::Cm$.

1.8.2. Construction of the modified strain *E. coli* MG1655 Δedd - Δeda $\Delta gloA$, $\Delta aldA$, $\Delta aldB$, $\Delta arcA::Km$, $\Delta ndh::Cm$

The deletion of the gene *ndh* by replacement of the gene by a chloramphenicol resistance cassette in the strain *E. coli* Δedd - Δeda $\Delta gloA$, $\Delta aldA$, $\Delta aldB$, $\Delta arcA::Km$, was performed by the technique of transduction with phage P1 (Protocol 2).

The deletion $\Delta ndh::Cm$ and the others modifications were checked using the oligonucleotides described in Table 6.

The strain obtained was named *E. coli* MG1655 Δedd - Δeda $\Delta gloA$, $\Delta aldA$, $\Delta aldB$, $\Delta arcA::Km$, $\Delta ndh::Cm$.

1.9. Construction of the modified strain *E. coli* MG1655 Δedd - Δeda , $\Delta gloA$, $\Delta aldA$, $\Delta aldB$, $\Delta arcA$, Δndh

The antibiotic resistance cassette was eliminated in the strain *E. coli* MG1655 Δedd - Δeda $\Delta gloA$, $\Delta aldA$, $\Delta aldB$, $\Delta arcA::Km$, $\Delta ndh::Cm$ according to Protocol 3.

The loss of the antibiotic resistance cassettes was checked by PCR analysis with the oligonucleotides given in Table 6. The presence of the modifications previously built in the strain was also checked using the oligonucleotides given in Table 6.

The strain obtained was named *E. coli* MG1655 Δedd - Δeda $\Delta gloA$, $\Delta aldA$, $\Delta aldB$, $\Delta arcA$, Δndh .

1.10. Construction of the modified strain *E. coli* MG1655 $\Delta gloA$, Δedd - Δeda , $\Delta aldA$, $\Delta aldB$, $\Delta arcA$, Δndh (pME101VB01-*yqhD**(G149E)-*gldA**(A160T))

1.10.1. Construction of the plasmid pME101-VB01-*yqhD**(G149E)-*gldA**(A160T)

1.10.1.1. Construction of the plasmid pME101-VB01

The plasmid pME101VB01 was derived from plasmid pME101 and harbors a multiple cloning site containing recognition site sequences specific for the rare restriction endonucleases *NheI*, *SnaBI*, *PacI*, *BglII*, *AvrII*, *SacII* and *AgeI* following by the *adc* transcription terminator of *Clostridium acetobutylicum* ATCC824.

For the expression from a low copy vector the plasmid pME101 was constructed as follows. The plasmid pCL1920 (Lerner & Inouye, 1990, NAR 18, 15 p 4631 - GenBank AX085428) was PCR amplified using the oligonucleotides PME101F and PME101R and the *Bst*Z17I-*Xmn*I fragment from the vector pTrc99A (Amersham Pharmacia Biotech, Piscataway, N.J) harboring the *lacI* gene and the *trc* promoter was inserted into the amplified vector.

PME101F (SEQ ID NO 12): ccgacagtaagacgggtaagcctg

PME101R (SEQ ID NO 13): agcttagtaaagccctcgctag

A synthetic double-stranded nucleic acid linker comprising the multicloning site and *adc* transcriptional terminator was used to generate pME101VB01. Two 100 bases oligonucleotides that complement flanked by *Nco*I or *Hind*III digested restriction sites were annealed. The 100-base pair product was subcloned into *Nco*I / *Hind*III digested plasmid pME101 to generate pME101VB01.

pME101VB01 1, consisting of 100 bases (SEQ ID NO 14):

catgggctagctacgtattaattaagatctcctagggagctcaccggtTAAAAATAAGAGTTACCTTAAAT
GGTAACTCTTATTTTTTTTAaggcgcgcca

pME101VB01 2, consisting of 100 bases (SEQ ID NO 15):

agcttggcgcgccTAAAAAAATAAGAGTTACCATTTAAGGTAACCTTATTTTTTAaccgg
tgagctccctaggagatctttaattaatacgtagctagcc

with:

- a region (underlined lower-case letters) corresponding to the multicloning site
- a region (upper-case letters) corresponding to the *adc* transcription terminator (sequence 179847 to 179814) of *Clostridium acetobutylicum* ATCC 824 pSOL1 (NC_001988).

1.10.1.2. Construction of the plasmid pME101-VB01-*yqhD**(G149E)-*gldA**(A160T)

1.10.1.2.1. Construction of the plasmid pSCB-*yqhD**(G149E)

The gene *yqhD* was PCR amplified from genomic DNA of *E. coli* MG1655 using the following oligonucleotides

yqhD F, consisting of 43 pb, (SEQ ID NO 16)

cgatgcacg**catga**acaactttaatctgcacaccccaaccg

with :

- a region (underlined letter) homologous to the sequence (3153377 to 3153408) of the gene *yqhD*
- a restriction site *Bsp*HI (bold face letters)

yqhD R, consisting of 29 pb (SEQ ID NO 17)

Ctag**ctagc**ttagcggcggttcgtata

With :

- a region (underlined letter) homologous to the sequence (3154540 to 3154521) the gene *yqhD*
- a restriction site *Nhe*I (bold face letters)

The PCR amplified fragment was cloned in pSCB (strataclone®). The resulting plasmid was named pSCB-*yqhD*. A directed mutagenesis was performed on this plasmid with the following oligonucleotides : *yqhD**G149EmutDirF (consisting of 45 pb, gggtcagaatccaacgcaga**a**agcggat**A**tcccgtaaaaccacaggc, (SEQ ID NO 18) and *yqhD**G149EmutDirR (consisting of 45 pb

gcctgtggttttacggga**T**atcaccgcttctgcgttggattctgaacc, (SEQ ID NO 19). The two oligonucleotides were homologous to the region 3153803 to 3153850. In bold face letter, bases which were changed to create the mutation G149E and capital letter, the base which were changed to create *EcoRV* restriction site. The resulted plasmid was named pSCB-
 5 *yqhD**(G149E).

1.10.1.2.2. Construction of the plasmid pSCB-*gldA**(A160T)

The gene *gldA*, PCR amplified from genomic DNA of *E. coli* MG1655 using the oligonucleotides *gldA* F and *gldA* R was cloned in pSCB (Strataclone®). The resulting plasmid was named pSCB-*gldA*.

10 A directed mutagenesis was performed on this plasmid with the following oligonucleotides: *gldA**A160TmutDirF (consisting of 45 pb, gacacaaaatcgctcgctggc**a**cacctgcacgtctgCtagcggcg, SEQ ID NO 20) and *gldA**A160TmutDirR (consisting of 45 pb cgccgctaGcagacgtgcaggtgtgccagcgacgattttggtgtc, SEQ ID NO 21). The two
 15 oligonucleotides are homologous to the region 4136602 to 4136558. In bold face letter, bases which were changed to create the mutation A160T and underlined letter, the base which were change to create *EcoRV* restriction site. The resulted plasmid was named pSCB-*gldA**(A160T).

1.10.1.3. Construction of the pME101VB01-*yqhD**(G149E)-*gldA**(A160T)

20 The pSCB-*yqhD**(G149E) was cut with the restriction enzymes *Bsp*HI and *Nhe*I and the fragment containing *yqhD**(G149E) was cloned into the *Nco*I/*Nhe*I sites of the vector pME101VB01. The resulting plasmid was named pME101VB01-*yqhD**(G149E). The pSCB-*gldA**(A160T) was cut with the restriction enzymes *avr*II and *Sac*I and the fragment containing *gldA**(A160T) was cloned into the *avr*II/*Sac*I sites of the vector
 25 pME101VB01-*yqhD**(G149E). The resulting plasmid was named pME101VB01-*yqhD**(G149S)-*gldA**(A160T).

1.10.2. Construction of the modified strain *E. coli* MG1655 Δ *gloA*, Δ *edd-eda*, Δ *aldA*, Δ *aldB*, Δ *arcA*, Δ *ndh* (pME101VB01-*yqhD**(G149E)-*gldA**(A160T))

30 The plasmid pME101VB01-*yqhD*-*gldA**(A160T) was introduced by electroporation into the strain *E. coli* MG1655 Δ *gloA*, Δ *edd-eda*, Δ *aldA*, Δ *aldB*, Δ *arcA*, Δ *ndh*.

The strain obtained was named *E. coli* MG1655 Δ *gloA*, Δ *edd-eda*, Δ *aldA*, Δ *aldB*, Δ *arcA*, Δ *ndh* (pME101VB01-*yqhD**(G149E)-*gldA**(A160T)).

35 **1.11. Construction of the modified strain *E. coli* MG1655 *mgsA**(H21Q)::Km Δ *edd-eda*, Δ *gloA*, Δ *aldA*, Δ *aldB*, Δ *arcA*, Δ *ndh* (pME101VB01-*yqhD**(G149E)-*gldA**(A160T))**

1.11.1. Construction of the modified strain *E. coli* MG1655 *mgsA**(H21Q)::Km

1.11.1.1. Construction of the modified strain *E. coli* MG1655 *mgsA**(H21Q)

5 A mutation was introduced in the *mgsA* gene in order to obtain the mutant protein MgsA*(H21Q). The technique used to build this modification was described by Heermann et al. (2008), Microbial Cell Factories.7(14): 1-8.

The following oligonucleotides were used to amplify the *rpsL*-Neo cassette :

1. *mgsA**(H21Q) ::*rpsL*-Neo F, consisting in 105 pb, (SEQ ID NO 22)
 10 gttaactacggatgtacattatggaactgacgactcgcaactttacctgcgcggaaacatattgcgctggtggcacacgatcaggcct
ggtgatgatggcgggatc

with,

- a region (underlined letter) homologous to the sequence of the gene *mgsA*.
- a region (bold face letter) to amplified *rpsL*-Neo cassette.

15 2. *mgsA**(H21Q) ::*rpsL*-Neo R (SEQ ID NO 23)
gggaaattaagtaccggtagtgccgtgtgcatacagtacgtgtgtccagtaacgggtgatgccgttcacccagctcat
cagcatctgtttgcattcagaagaactcgtaagaagg

with,

- a region (underlined letter) homologous to the sequence of the gene *mgsA* with
- 20 two mutation, the first one (in red) to create the mutation H21Q and the second one (in yellow) to create the restriction site *AlwN1*.
- a region (bold face letter) to amplify *rpsL*-Neo cassette.

The fragment obtained was introduced into the strain MG1655 *rpsL** (built as described in Heermann et al.) according to Protocol 1. The strain obtained was checked by PCR and
 25 sequence analysis. The strain obtained is named *E. coli mgsA**(H21Q)::*rpsL*-Neo.

The deletion of the cassette *rpsL*-Neo was performed according to Protocol 1. The fragment transformed was obtained by the restriction with *NcoI* and *SacI* of the plasmid pETTOPO-*mgsA**(H21Q).

The modification was checked by PCR using oligonucleotides described in Table 6.

30 The strain obtained was named strain *E. coli* MG1655 *mgsA**(H21Q).

1.11.1.2. Construction of the modified strain *E. coli* MG1655 *mgsA**(H21Q)::Km

A kanamycin resistance cassette was introduced in 3' of *mgsA**(H21Q) open reading frame (ORF) using the following primers :

35 *mgsA* ::Km F consisting of 100 bp : (SEQ ID NO 24)
tccagtcgccgcattcaacgacgcggtcgatattctgatccccgattatcagcggtatctcgccgaccgtctgaagtaatgtaggct
ggagctgcttcg

with :

- a region (underlined letters) homologous to the end of *mgsA**(H21Q) ORF ,
- a region (Bold letter) to amplified Kanamycin cassette.

mgsA ::Km R consisting of 100 bp : (SEQ ID NO 25)

Tgtggaaatactgaaaaatctggatgtgccggtggcgagaaaaccgtaagaaacaggtggcgtttggcacctgtgcaata**catat**
5 **gaatatectccttag**

- a region (underlined letters) homologous to the end of *helD* ORF ,
- a region (Bold letter) to amplified Kanamycin cassette.

The fragment obtained was introduced into the strain MG1655 *mgsA**(H21Q) according to Protocol 1. The strain obtained was checked by PCR. The strain obtained was named *E.*

10 *coli mgsA**(H21Q)::Km

1.11.2. Construction of the modified strain *E. coli* MG1655 *mgsA(H21Q)::Km Δ *gloA*, Δ *edd-eda*, Δ *aldA*, Δ *aldB*, Δ *arcA*, Δ *ndh* (*pME101VB01-yqhD**(G149E)-*gldA**(A160T))**

15 The replacement of the *mgsA* with the *mgsA**(H21Q)::Km into the strain *E. coli* Δ *gloA*, Δ *edd-eda*, Δ *aldA*, Δ *aldB*, Δ *arcA*, Δ *ndh* (*pME101VB01-yqhD**(G149E)-*gldA**(A160T)) was performed by the technique of transduction with phage P1. IPTG was added to the culture to promote the expression of the genes borne on the plasmid.

The modifications *mgsA**(H21Q)::Km and the others deletion were checked using the
20 oligonucleotides described in Table 6.

2. Construction of the modified strain *E. coli* MG1655 *mgsA*::Km, Δ *edd-eda*, Δ *gloA*, Δ *aldA*, Δ *aldB*, Δ *arcA*, Δ *ndh* (*pME101-VB01-yqhD(G149E)-*gldA**(160T)).**

25 **2.1. Construction of the modified strain *E. coli* MG1655 Δ *gloA*, Δ *edd-eda*, Δ *aldA*, Δ *aldB*, Δ *arcA*, Δ *ndh* (*pME101VB01-yqhD**(G149E)-*gldA**(A160T))**

The construction of this strain is described above.

2.2. Construction of the modified strain *E. coli* MG1655 *mgsA*::Km Δ *edd-eda*, Δ *gloA*, Δ *aldA*, Δ *aldB*, Δ *arcA*, Δ *ndh* (*pME101VB01-yqhD(G149E)-*gldA**(A160T))**

2.2.1. Construction of the modified strain *E. coli* MG1655 *mgsA*::Km

A kanamycin resistance cassette was introduced in 3' of *mgsA* open reading frame using the following primers : as shown in SEQ ID NO 24 and SEQ ID NO 25.

The fragment obtained was introduced into the strain MG1655 according to Protocol 1.

35 The strain obtained was checked by PCR. The strain obtained was named *E. coli mgsA*::Km

2.2.2 Construction of the modified strain *E. coli* MG1655 mgsA::Km, Δ gloA,

Δ edd-eda, Δ aldA, Δ aldB, Δ arcA, Δ ndh (pME101VB01-yqhD*(G149E)-gldA*(A160T))

The replacement of the mgsA with the mgsA::Km into the strain *E. coli* Δ edd-eda Δ gloA, Δ aldA, Δ aldB, Δ arcA, Δ ndh (pME101VB01-yqhD*(G149E)-gldA*(A160T)) was performed by the technique of transduction with phage P1. The modifications mgsA::Km and the others deletion were checked using the oligonucleotides described in Table 6.

The resulting strain was named *E. coli* MG1655 mgsA::Km, Δ gloA, Δ edd-eda, Δ aldA, Δ aldB, Δ arcA, Δ ndh (pME101VB01-yqhD*(G149E)-gldA*(A160T)).

3. Assessment of 1,2-propanediol production in two *E. coli* isogenic strains differing only in the mgsA alleles

The two strains described above were cultivated in an Erlenmeyer flask assay (500 ml flasks with 50 ml of medium) under aerobic conditions in minimal medium MML11PG1_100 (see composition in Table 9) with 20 g/l glucose as sole carbon source. Spectinomycin was added at a concentration of 50 mg/l.

Table 9 : composition of minimal medium MML11PG1_100.

Constituent	Concentration (g/l)
EDTA	0.0084
CoCl ₂ 6H ₂ O	0.0025
MnCl ₂ 4H ₂ O	0.0150
CuCl ₂ 2H ₂ O	0.0015
H ₃ BO ₃	0.0030
Na ₂ MoO ₄ 2H ₂ O	0.0025
Zn(CH ₃ COO) ₂ 2H ₂ O	0.0130
Fe(III) citrate H ₂ O	0.1064
Citric acid	1.70
KH ₂ PO ₄	1.65
K ₂ HPO ₄ 3H ₂ O	0.92
(NH ₄) ₂ HPO ₄	0.40
(NH ₄) ₂ SO ₄	4.88
MgSO ₄ 7H ₂ O	1.00
CaCl ₂ 2H ₂ O	0.08
Thiamine	0.01
Glucose or Sucrose	20.00
MOPS buffer	40.00

The pH of the medium was adjusted to 6.8 with sodium hydroxide

The culture was carried out at 37°C and the pH was maintained by buffering the culture medium with MOPS.

At the end of the culture, 1,2-propanediol and residual glucose in the fermentation broth were analysed by HPLC using a Biorad HPX 97H column for the separation and a

refractometer for the detection. The yields of 1,2-propanediol over glucose were then calculated.

Table 10: production of 1,2-propanediol in minimal medium with glucose as carbon source.

Strain /	Carbon source	1,2-propanediol titer (g/l)	1,2-propanediol yield (g/g carbon source)
<i>E. coli</i> MG1655 <i>mgsA</i> ::Km Δ <i>gloA</i> Δ <i>edd-eda</i> Δ <i>aldA</i> Δ <i>aldB</i> Δ <i>arcA</i> Δ <i>ndh</i> (pME101VB01- <i>yqhD</i> *(G149E)- <i>gldA</i> *(A160T))	glucose	ND (n=3)	- (n=3)
<i>E. coli</i> MG1655 <i>mgsA</i> *(H21Q)::Km Δ <i>gloA</i> Δ <i>edd-eda</i> Δ <i>aldA</i> Δ <i>aldB</i> Δ <i>arcA</i> Δ <i>ndh</i> (pME101VB01- <i>yqhD</i> *(G149E)- <i>gldA</i> *(A160T))	glucose	0.904 +/- 0.049 (n=3)	0.133 +/- 0.009 (n=3)

5 ND means 'not detected' - n is the number of repetitions of the same experiment - The figures given are the mean and standard deviation of the figures obtained for n repetitions.

The strain with wild-type MGS does not produce any 1,2-propanediol due to the inhibition by inorganic phosphate.

10 The *E. coli* strain with a mutant MGS clearly produces 1,2-propanediol, confirming together with the above findings that strains with mutant MGS that are insensitive to inhibition by inorganic phosphate can produce 1,2-propanediol in the presence of inorganic phosphate.

15 **Example 4 : Production of 1,2-propanediol by *E. coli* with a mutant MGS, a mutant YqhD and a mutant GlyDH on glucose and sucrose**

1. Construction of the modified strain *E. coli* MG1655 *mgsA(H21Q)::Km Δ *gloA*, Δ *edd-eda*, Δ *aldA*, Δ *aldB*, Δ *arcA*, Δ *ndh* (pME101-VB01-*yqhD**(G149E)-*gldA**(A160T))**

The construction of this strain was described previously.

20

2. Construction of the modified strain *E. coli* MG1655 *mgsA(H21Q) Δ *gloA*, Δ *edd-eda*, Δ *aldA*, Δ *aldB*, Δ *arcA*, Δ *ndh* (pME101-VB01-*yqhD**(G149E)-*gldA**(A160T)) (pBBR1MCS5-*cscBKAR*)**

2.1. Construction of the plasmid pBBR1MCS5-*cscBKAR*

25 The plasmid pKJL101.1 (Jahreis et al. (2002), J. Bacteriol. 184:5307-5316) was digested by *EcoRI*. The fragment containing the *cscBKAR* gene was cloned in pBBR1MCS5 (Kovach et al. (1995), Gene, 166 175-176) also digested by *EcoRI*.

The resulted plasmid was named pBBR1MCS5-*cscBKAR*.

2.2. Construction of the modified strain *E. coli* MG1655 *mgsA(H21Q)::Km, Δ *gloA*, Δ *edd-eda*, Δ *aldA*, Δ *aldB*, Δ *arcA*, Δ *ndh* (pME101-VB01-*yqhD**(G149E)-*gldA**(A160T)) (pBBR1MCS5-*cscBKAR*)**

5 The plasmids pME101-VB01-*yqhD**(G149E)-*gldA**(A160T) and pBBR1MCS5-*cscBKAR* were introduced by electroporation in the strain *E. coli* MG1655 *mgsA**(H21Q), Δ *edd-eda*, Δ *gloA*, Δ *aldA*, Δ *aldB*, Δ *arcA*, Δ *ndh*.

The strain obtained was named strain *E. coli* MG1655 *mgsA**(H21Q)::Km, Δ *gloA*, Δ *edd-eda*, Δ *aldA*, Δ *aldB*, Δ *arcA*, Δ *ndh* (pME101-VB01-*yqhD**(G149E)-*gldA**(A160T))
10 (pBBR1MCS5-*cscBKAR*).

3. Assessment of 1,2-propanediol production in two *E. coli* with a mutant MGS, a mutant YqhD and a mutant GlyDH on glucose and sucrose

15 The two strains described above were cultivated in an Erlenmeyer flask assay (500 ml flasks with 50 ml of medium) under aerobic conditions in minimal medium MML11PG1_100 (see composition in Table 9) with 20 g/l glucose or sucrose as sole carbon source. Spectinomycin was added at a concentration of 50 mg/l.

20 The culture was carried out at 37°C and the pH was maintained by buffering the culture medium with MOPS.

At the end of the culture, 1,2-propanediol and residual glucose or sucrose in the fermentation broth were analysed by HPLC using a Biorad HPX 97H column for the separation and a refractometer for the detection. The yields of 1,2-propanediol over glucose or sucrose were then calculated.

25 Table 11: production of 1,2-propanediol in minimal medium with glucose or sucrose as carbon source.

Strain	Carbon source	1,2-propanediol titer (g/l)	1,2-propanediol yield (g/g carbon source)
<i>E. coli</i> MG1655 <i>mgsA</i> *(H21Q)::Km Δ <i>gloA</i> Δ <i>edd-eda</i> Δ <i>aldA</i> Δ <i>aldB</i> Δ <i>arcA</i> Δ <i>ndh</i> (pME101VB01- <i>yqhD</i> *(G149E)- <i>gldA</i> *(A160T))	glucose	0.904 +/- 0.049 (n=3)	0.133 +/- 0.009 (n=3)
<i>E. coli</i> MG1655 <i>mgsA</i> *(H21Q)::Km Δ <i>gloA</i> Δ <i>edd-eda</i> Δ <i>aldA</i> Δ <i>aldB</i> Δ <i>arcA</i> Δ <i>ndh</i> (pME101VB01- <i>yqhD</i> *(G149E)- <i>gldA</i> *(A160T)) (pBBR1MCS5- <i>cscBKAR</i>)	sucrose	1.823 +/- 0.098 (n=3)	0.196 +/- 0.007 (n=3)

n is the number of repetitions of the same experiment - The figures given are the mean and standard deviation of the figures obtained for n repetitions.

The production of 1,2-propanediol in a *E. coli* strain with a mutant MGS was improved on sucrose as sole carbon source as compared with glucose.

Example 5 : Construction of two *E. coli* acetol producer strains expressing wildtype or modified MGS and assessment of acetol production

1. Construction of the modified strain *E. coli* MG1655 *mgsA(H21Q)::Km *P_{trc01}-gapA*::cm, Δ *edd-eda*, Δ *gloA*, Δ *aldA*, Δ *aldB*, Δ *gldA*, (pME101-VB01-*yqhD*)**

1.1. Construction of the modified strain *E. coli* MG1655 Δ *gloA*, Δ *edd-eda* Δ *aldA*, Δ *aldB*, Δ *gldA*

1.1.1. Construction of the modified strain Δ *gldA*::Km

The gene *gldA* was inactivated in strain *E. coli* MG1655 by inserting a Kanamycin antibiotic cassette and deleting most of the gene concerned using the technique described in Protocol 1 with the oligonucleotides given in Table 5. The deletion was checked by a PCR analysis with the appropriate oligonucleotides given in Table 5.

The resulting strain was named *E. coli* MG1655 Δ *gldA*::Km.

1.1.2. Construction of the modified strain *E. coli* MG1655 Δ *edd-eda* Δ *gloA*, Δ *aldA*, Δ *aldB*, Δ *gldA*::Km

The deletion of the gene *gldA* by replacement of the gene by a Kanamycine resistance cassette in the strain *E. coli* MG1655 Δ *edd-eda* Δ *gloA*, Δ *aldA*, Δ *aldB* was performed by the technique of transduction with phage P1 (Protocol 2).

The deletion was checked by a PCR analysis with the appropriate oligonucleotides given in Table 6 as well as the other deletions already present in the strain.

The resulting strain was named *E. coli* MG1655 Δ *edd-eda* Δ *gloA*, Δ *aldA*, Δ *aldB*, Δ *gldA*::Km

1.1.3. Construction of the modified strain *E. coli* MG1655 Δ *edd-eda* Δ *gloA*, Δ *aldA*, Δ *aldB*, Δ *gldA*

The antibiotic resistance cassette was eliminated in the strain *E. coli* MG1655 Δ *edd-eda* Δ *gloA*, Δ *aldA*, Δ *aldB*, Δ *gldA*::Km according to Protocol 3.

The loss of the antibiotic resistance cassette was checked by PCR analysis with the oligonucleotides given in Table 6. The presence of the modifications previously built in the strain was also checked using the oligonucleotides given in Table 6.

The strain obtained was named *E. coli* MG1655 Δ *edd-eda* Δ *gloA*, Δ *aldA*, Δ *aldB*, Δ *gldA*.

1.2. Construction of the modified strain *E. coli* MG1655 Ptrc01-gapA::cm Δ gloA, Δ edd-eda, Δ aldA, Δ aldB, Δ gldA

1.2.1. Construction of the modified strain MG1655 Ptrc01-gapA::Cm

The replacement of the natural gapA promoter with the synthetic short Ptrc01 promoter (SEQ ID NO 26 : gagctgttgactattaatcatccggctcgaataatgtgtgg) into the strain *E. coli* MG1655 was made by replacing 225 pb of upstream gapA sequence with FRT-CmR-FRT and an engineered promoter using the technique described in Protocol 2 with the oligonucleotides given in Table 5.

The modification was checked by a PCR analysis with the appropriate oligonucleotides given in Table 6. The resulting strain was named *E. coli* MG1655 Ptrc01-gapA::Cm.

1.2.2. Construction of the modified strain *E. coli* MG1655 Ptrc01-gapA::cm Δ edd-eda Δ gloA, Δ aldA, Δ aldB, Δ gldA

The replacement of the natural gapA promoter with the synthetic short Ptrc01 promoter into the strain *E. coli* MG1655 Δ edd-eda Δ gloA, Δ aldA, Δ aldB, Δ gldA was performed by the technique of transduction with phage P1.

The modification Ptrc01-gapA::cm and the other deletions were checked using the oligonucleotides described in Table 6.

The strain obtained was named *E. coli* MG1655 Ptrc01-gapA::cm Δ edd-eda Δ gloA, Δ aldA, Δ aldB, Δ gldA.

1.3. Construction of the modified strain *E. coli* MG1655 Ptrc01-gapA::cm Δ edd-eda Δ gloA, Δ aldA, Δ aldB, Δ gldA pME101-VB01-yqhD*(G149E),

The plasmid pME101VB01-yqhD*(G149E) was introduced by electroporation into the strain *E. coli* MG1655 Ptrc01-gapA::cm Δ edd-eda Δ gloA, Δ aldA, Δ aldB, Δ gldA.

The strain obtained was named *E. coli* MG1655 Ptrc01-gapA::cm, Δ gloA, Δ edd-eda, Δ aldA, Δ aldB, Δ gldA (pME101-VB01-yqhD*(G149E)).

1.4. Construction of the modified strain *E. coli* MG1655 mgsA*(H21Q)::Km Ptrc01-gapA::cm Δ edd-eda Δ gloA, Δ aldA, Δ aldB, Δ gldA pME101-VB01-yqhD*(G149E)

The replacement of the mgsA with the mgsA*(H21Q)::Km into the strain *E. coli* Δ edd-eda Δ gloA, Δ aldA, Δ gldA (pME101VB01-yqhD*(G149E)) was performed by the technique of transduction with phage P1. IPTG was added to the culture to promote the expression of the gene borne on the plasmid.

The modification mgsA*(H21Q)::Km and the other deletions were checked using the oligonucleotides described in Table 6.

The resulting strain was named *E. coli* MG1655 mgsA*(H21Q)::Km Ptrc01-gapA::cm Δ edd-eda Δ gloA, Δ aldA, Δ aldB, Δ gldA pME101-VB01-yqhD*(G149E).

2. Construction of the modified strain *E. coli* MG1655 *P_{trc01}-gapA::cm*, *mgsA::Km* Δ *edd-eda*, Δ *gloA*, Δ *aldA*, Δ *aldB*, Δ *gldA*, (pME101-VB01-*yqhD**(G149E))

2.1. Construction of the modified strain *E. coli* MG1655 *P_{trc01}-gapA::cm*, *mgsA::Km* Δ *edd-eda*, Δ *gloA*, Δ *aldA*, Δ *aldB*, Δ *gldA* (pME101-VB01-*yqhD**(G149E))

The *mgsA::Km* was constructed in the strain *E. coli* MG1655 *P_{trc01}-gapA::cm* Δ *edd-eda*, Δ *gloA*, Δ *aldA*, Δ *aldB*, Δ *gldA* (pME101-VB01-*yqhD**(G149E)) as previously described.

The strain obtained was named strain *E. coli* MG1655 *P_{trc01}-gapA::cm*, *mgsA::Km*, Δ *edd-eda*, Δ *gloA*, Δ *aldA*, Δ *aldB*, Δ *gldA*. (pME101-VB01-*yqhD**(G149E))

3. Assessment of acetol production in two *E. coli* isogenic strains differing only in the *mgsA* alleles

The two strains described above were cultivated in an Erlenmeyer flask assay (500 ml flasks with 50 ml of medium) under aerobic conditions in minimal medium MML11PG1_100 (see composition in Table 9) with 20 g/l glucose as sole carbon source. Spectinomycin was added at a concentration of 50 mg/l.

The culture was carried out at 37°C and the pH was maintained by buffering the culture medium with MOPS.

At the end of the culture, acetol and residual glucose in the fermentation broth were analysed by HPLC using a Biorad HPX 97H column for the separation and a refractometer for the detection. The yields of acetol over glucose or sucrose were then calculated.

Table 12: production of acetol in minimal medium with glucose as carbon source.

Strain	Carbon source	acetol titer (g/l)	acetol yield (g/g carbon source)
<i>E. coli</i> MG1655 <i>mgsA::Km</i> Δ <i>gloA</i> Δ <i>edd-eda</i> Δ <i>aldA</i> Δ <i>aldB</i> Δ <i>gldA</i> (pME101VB01- <i>yqhD</i> *(G149E))	glucose	0.27 +/- 0.37 (n=2)	0.014 +/- 0.020 (n=2)
<i>E. coli</i> MG1655 <i>mgsA</i> *(H21Q):: <i>Km</i> Δ <i>gloA</i> Δ <i>edd-eda</i> Δ <i>aldA</i> Δ <i>aldB</i> Δ <i>gldA</i> (pME101VB01- <i>yqhD</i> *(G149E))	glucose	2.23 +/- 0.23 (n=2)	0.272 +/- 0.022 (n=2)

n is the number of cultures of different clones of the same strain - The figures given are the mean and standard deviation of the figures obtained for n cultures.

The production of acetol in the *E. coli* strain with a mutant MGS was dramatically improved as compared with the isogenic strain with a native MGS.

Example 6 : Production of acetol by *E. coli* with a mutant MGS and a mutant YqhD on glucose and sucrose

1. Construction of the modified strain *E. coli* MG1655 Ptrc01-gapA::cm,
5 mgsA*(H21Q::Km), Δ edd-eda Δ gloA, Δ aldA, Δ aldB, Δ gldA, pME101-VB01-yqhD*(G149E),

The construction of this strain was described previously.

2. Construction of the modified strain *E. coli* MG1655 Ptrc01-gapA::cm,
10 mgsA*(H21Q), Δ edd-eda Δ gloA, Δ aldA, Δ aldB, Δ gldA pME101-VB01-yqhD*(G149E)
pBBR1MCS5-cscBKAR

The plasmid pBBR1MCS5-cscBKA was introduced by electroporation in the strain *E. coli* MG1655 Ptrc01-gapA::cm, mgsA*(H21Q::Km), Δ edd-eda Δ gloA, Δ aldA, Δ aldB, Δ gldA pME101-VB01-yqhD*(G149E)

- 15 The resulted strain was named *E. coli* MG1655 Ptrc01-gapA::cm, mgsA*(H21Q)::Km, Δ edd-eda Δ gloA, Δ aldA, Δ aldB, Δ gldA pME101-VB01-yqhD*(G149E) pBBR1MCS5-cscBKAR.

Example 7 : Construction of two *E. coli* lactate producer strains expressing wildtype or modified MGS and assessment of lactate production

- 1 - Construction of a modified strain *E. coli* MG1655 Ptrc01-gapA, Δ edd-eda, Δ yqhD, Δ lld, Δ lldD (pJB137-PgapA-ppsA) (pME101-VB01-yedU).

- 1.1. Construction of the modified strain *E. coli* MG1655 Ptrc01-gapA, Δ edd-eda
25 1.1.1. Construction of the strain *E. coli* MG1655 Ptrc01-gapA::cm, Δ edd-eda::Km

The deletion of the gene *edd-eda* by replacement of the gene by a kanamycin resistance cassette in the strain *E. coli* MG1655 Ptrc01-gapA::cm (see Example 3) is performed by the technique of transduction with phage P1 (protocole 2).

- 30 The deletion is checked by a PCR analysis with the appropriate oligonucleotides given in table 6 as well as the others deletions present in the strain.

The resulting strain is named *E. coli* MG1655 Ptrc01-gapA::cm, Δ edd-eda::Km

- 1.1.2. Construction of the strain *E. coli* MG1655 Ptrc01-gapA, Δ edd-eda
The antibiotic resistance cassette is eliminated in the strain *E. coli* MG1655 Δ Ptrc01-gapA::cm, Δ edd-eda::Km according to protocol 3.
35

The loss of the antibiotic resistance cassettes is checked by PCR analysis with the oligonucleotides given in Table 6. The presence of the modifications previously built in the strain is also checked using the oligonucleotides given in Table 6.

The strain obtained is named *E. coli* MG1655 Ptrc01-gapA, Δ edd-eda.

1.2. Construction of the *E. coli* MG1655 Ptrc01-gapA, Δ edd-eda, Δ yqhD.

1.2.1. Construction of the modified strain *E. coli* MG1655 Δ yqhD::Km

5 The gene *yqhD* is inactivated by inserting a kanamycin antibiotic resistance cassette and deleting most of the gene concerned using the technique described in Protocol 1 with the oligonucleotides given in Table 5. The deletion is checked by a PCR analysis with the appropriate oligonucleotides given in table 6.

The strain obtained is named *E. coli* MG1655 Δ yqhD::Km.

10 1.2.2. Construction of the modified strain *E. coli* Ptrc01-gapA, Δ edd-eda, Δ yqhD

The deletion of the gene *yqhD* by replacement of the gene by a Kanamycin resistance cassette in the strain *E. coli* MG1655 Ptrc01-gapA Δ edd-eda is performed using the transduction technique with phage P1 described in Protocol 2. The deletion is checked
15 by a PCR analysis with the appropriate oligonucleotides given in table 6.

The strain obtained is named *E. coli* Ptrc01-gapA, Δ edd-eda, Δ yqhD::Km..

1.3. Construction of the modified strain *E. coli* Ptrc01-gapA, Δ edd-eda, Δ yqhD, Δ dld

20 1.3.1. Construction of the modified strain *E. coli* MG1655 Δ dld::Cm

The gene *dld* is inactivated by inserting a chloramphenicol antibiotic resistance cassette and deleting most of the gene concerned using the technique described in Protocol 2 with the oligonucleotides given in Table 5. The deletion is checked by a PCR analysis with the appropriate oligonucleotides given in table 6.

25 The strain obtained is named *E. coli* MG1655 Δ dld::Cm.

1.3.2. Construction of the modified strain *E. coli* MG1655 Ptrc01-gapA, Δ edd-eda, Δ yqhD, Δ dld.

The deletion of the gene *dld* in the strain *E. coli* MG1655 *E. coli* MG1655 Ptrc01-gapA, Δ edd-eda, Δ yqhD::Km is performed using the transduction technique with phage P1
30 described in Protocol 2.

The resulting strain is named *E. coli* MG1655 Ptrc01-gapA, Δ edd-eda, Δ yqhD::Km, Δ dld::Cm.

The chloramphenicol and kanamycin resistance cassettes are then eliminated according to Protocol 3.

35 The strain obtained is named *E. coli* MG1655 Ptrc01-gapA, Δ edd-eda, Δ yqhD, Δ dld.

1.4. Construction of the modified strain *E. coli* Ptrc01-gapA, Δ edd-eda, Δ yqhD, Δ dld, Δ lldD

1.4.1. *E. coli* MG1655 Δ lldD::Cm

The gene *lldD* is inactivated by inserting a chloramphenicol antibiotic resistance cassette and deleting most of the gene concerned using the technique described in Protocol 2 with the oligonucleotides given in Table 5. The deletion is checked by a PCR analysis with the appropriate oligonucleotides given in Table 6.

The strain obtained is named *E. coli* MG1655 Δ lldD::Cm.

1.4.2. *E. coli* MG1655 MG1655 Ptrc01-gapA, Δ edd-eda, Δ yqhD, Δ dld, Δ lldD

The deletion of the gene *lldD* in the strain *E. coli* MG1655 Ptrc01-gapA, Δ edd-eda, Δ yqhD, Δ dld is performed using the transduction technique with phage P1 described in Protocol 3.

The resulting strain is named *E. coli* MG1655 Ptrc01-gapA, Δ edd-eda, Δ yqhD, Δ dld, Δ lldD::Cm.

The chloramphenicol resistance cassette is then eliminated according to Protocol 3. The strain obtained is named *E. coli* MG1655 Ptrc01-gapA, Δ edd-eda, Δ yqhD, Δ dld, Δ lldD.

1.5. Construction of the modified strain *E. coli* MG1655 Ptrc01-gapA, Δ edd-eda, Δ yqhD, Δ dld, Δ lldD (pJB137-PgapA-ppsA) (pME101-VB01-yedU)

The plasmids pJB137-PgapA-ppsA and pME101-VB01-yedU (described in patent application PCT/EP2009/053093) are introduced by electroporation in the strain *E. coli* MG1655 Ptrc01-gapA, Δ edd-eda, Δ yqhD, Δ dld, Δ lldD.

The resulted strain is named *E. coli* MG1655 Ptrc01-gapA, Δ edd-eda, Δ yqhD, Δ dld, Δ lldD. (pJB137-PgapA-ppsA) (pME101-VB01-yedU).

2. - Construction of a modified strain *E. coli* MG1655 Ptrc01-gapA, *mgsA**(H21Q), Δ edd-eda, Δ yqhD, Δ dld, Δ lldD (pJB137-PgapA-ppsA) (pME101-VB01-yedU)

2.1. Construction of a modified strain *E. coli* MG1655 Ptrc01-gapA, *mgsA**(H21Q), Δ edd-eda, Δ yqhD, Δ dld, Δ lldD

The mutation *mgsA**(H21Q) is constructed in the strain *E. coli* MG1655 Ptrc01-gapA, Δ edd-eda, Δ yqhD, Δ dld, Δ lldD as previously described in Example 3.

The strain obtained is named strain *E. coli* MG1655 Ptrc01-gapA, *mgsA**(H21Q), Δ edd-eda, Δ yqhD, Δ dld, Δ lldD.

2.2. Construction of a modified strain *E. coli* MG1655 Ptrc01-gapA, mgsA*(H21Q), Δ edd-eda, Δ yqhD, Δ dld, Δ lldD (pJB137-PgapA-ppsA) (pME101-VB01-yedU)

The plasmids pJB137-PgapA-ppsA and pME101-VB01-yedU, are introduced by electroporation in the strain *E. coli* MG1655 Ptrc01-gapA, mgsA*(H21Q), Δ edd-eda, Δ yqhD, Δ dld, Δ lldD.

The resulted strain is named *E. coli* MG1655 Ptrc01-gapA, mgsA*(H21Q), Δ edd-eda, Δ yqhD, Δ dld, Δ lldD (pJB137-PgapA-ppsA) (pME101-VB01-yedU).

10 **Example 8 : Construction of two *Saccharomyces cerevisiae* 1,2-propanediol producer strains and assessment of 1,2-propanediol production**

1 -Construction of two *S. cerevisiae* strains CENPK Δ gpd2, Δ tpi1, gldA*(A160T), yqhD, mgsA*(H21Q) and CENPK Δ gpd2, Δ tpi1, gldA*(A160T), yqhD*(G149E), mgsA*(H21Q)

1-1. Construction of *S. cerevisiae* strain CENPK Δ gpd2, gldA*(A160T).

The *S. cerevisiae* strain used was CEN.PK2-1C (*MATa*; *ura3-52*; *trp1-289*; *leu2-3,112*; *his3 Δ 1*; *MAL2-8C*; *SUC2*) from Euroscarf.

20 The gene *GPD2* was inactivated by transforming the strain CEN.PK2-1C with a PCR fragment corresponding to pTDH3-gldA*(A160T)-CYCt-pTEF1-ble-TEF1t cassette, built using the “short flanking homology” (SFH) method described by Guldener et al. (1996).

The pTDH3-gldA*(A160T)-CYCt-pTEF1-ble-TEF1t cassette was constructed using long PCR-based fusion of several fragments as described by Shevchuk et al. (2004).

25 pTDH3 and CYCt were amplified from the plasmid p406TDH3 (Addgene) using pTDH3/GPD2 F and pTDH3 R primers and CYCt /gldA F and CYCt/ Zeo R primers respectively. gldA*(A160T)- was amplified from pSCB gldA*(A160T)- using primers gldA / TDH3F and gldA /CYCtR.

30 pTEF1-ble-TEF1t was amplified from the plasmid pUG66 from Euroscarf using Zeo/CYCt F and ZEO/GPD2 R as primers.

All fragments were amplified using primers having overlapping ends as described in Table 13. Each fragment was then purified.

100 ng of each fragment was used in a PCR experiment without primers, using low annealing conditions allowing their simultaneous fusion.

35 The unpurified product obtained in this step was used as a matrix a in a PCR experiment at high *T*_m, using pTDH3/ GPD2 F and ZEO/GPD2 R primers having an extension of 40 bp homologous to the 40 first and 40 last bp of the *GPD2* locus (Table 13).

This fragments was integrated in the *GPD2* locus, replacing the *GPD2* open reading frame.

The transformation method used was the lithium acetate method described by Schiestl and Gietz (1989). The strain CENPK, $\Delta gpd2$, *gldA**(A160T) was selected on YEPD rich medium (1% bacto yeast extract, 2% bacto peptone, 2% glucose) supplemented with 75 μ g/ml of phleomycin (Cayla, France). The integration of *gldA**(A160T) and the deletion of *GPD2* gene were confirmed by PCR on genomic DNA extracted, using GPD2 ver F and GPD2 ver R primers (Table13).

This resulted in the heterologous expression of *gldA**(A160T) and deletion of *GPD2*. The resulting strain was named CENPK $\Delta gpd2$, *gldA**(A160T).

Table 13

Primer Name	Sequence	Description
pTDH3/ GPD2 F	<u>ATG CTT GCT GTC AGA AGA TTA</u> <u>ACA AGA TAC ACA TTC CTT AGT</u> TTA TCA TTA TCA ATA CTC G (SEQ ID NO 27)	The underlined 40 nucleotides correspond to the 40 first bp of <i>GPD2</i> gene 21 nucleotides in bold are homologous to the 21 first bp of pTDH3
pTDH3 R	ATCCTCGAACTAAGTTCTTGGT (SEQ ID NO 28)	23 nucleotides homologous to the 23 last bp of pTDH3
<i>gldA</i> / TDH3F	<u>AAA CAC CAG AAC TTA GTT TCG</u> <u>AAC TAG TTT ATT CCC ACT CTT</u> (SEQ ID NO 29)	The underlined 22 nucleotides underlined correspond to the last bp of pTDH3 20 nucleotides in bold are homologous to the 20 first bp of <i>gldA</i> or <i>gldA</i> * (A160T)
<i>gldA</i> /CYCtR	<u>TGA AAT ATA AAT AAC GTT CTT</u> <u>AAT ACT AAC ATA ACT ATA AAC</u> TAG TAT GGA CCG CAT TAT TC (SEQ ID NO 30)	The underlined 42 nucleotides correspond to the 42 first bp of CYCt 20 nucleotides in bold are homologous to the 20 last bp of <i>gldA</i> or <i>gldA</i> *(A160T)
CYCt / <i>gldA</i> F	<u>CCT GGA TGT ATT TAC CCG GTG</u> <u>ATT GAA TAA TGC GGT CCA TAC</u> <u>TAG TTT ATA GTT ATG TTA GTA</u> TTA (SEQ ID NO 31)	The underlined 46 nucleotides correspond to the 46 last bp of <i>gldA</i> or <i>gldA</i> 20 nucleotides in bold are homologous to the 20 first bp of CYCt
CYCt/ Zeo R	<u>GAG GCA AGC TAA ACA GAT CTC</u> <u>TAG ACC TAG GTA CCC GCC GGC</u> AAA TTA AAG CCT TCG AGC (SEQ ID NO 32)	The underlined 40 nucleotides correspond to the 40 first bp of <i>ble</i> gene 20 nucleotides in bold are homologous to the 20 last bp of CYCt
Zeo /CYCt F	<u>GCT TGA GAA GGT TTT GGG ACG</u> <u>CTC GAA GGC TTT AAT TTG CTA</u> GGT CTA GAG ATC TGT TTA GC (SEQ ID NO 33)	The underlined 40 nucleotides correspond to the 40 last bp of CYCt 22 nucleotides in bold are homologous to the 22 first bp of TEF1p

ZEO/GPD2 R	<u>CTA TTC GTC ATC GAT GTC TAG</u> <u>CTC TTC AAT CAT CTC CGG TCC</u> ACT AGT GGA TCT GAT ATC ACC T (SEQ ID NO 34)	The underlined 40 nucleotides correspond to the 40 last bp of <i>GPD2</i> gene 24 nucleotides in bold are homologous to the 24 last bp of <i>TEF1t</i>
GPD2 ver F	ATG CTT GCT GTC AGA AGA TT (SEQ ID NO 35)	20 nucleotides homologous to the 20 first bp of <i>GPD2</i> gene
GPD2 ver R	TAG TAT GGA CCG CAT TAT TC (SEQ ID NO 36)	20 nucleotides homologous to the 20 last nucleotides of <i>gldA*</i> (A160T)

1-2. Construction of two *S. cerevisiae* strains CENPK $\Delta gpd2$, *gldA**(A160T), *yqhD* and CENPK $\Delta gpd2$, *gldA**(A160T), *yqhD**(G149E)

- 5 The strain used was CENPK, $\Delta gpd2$, *gldA**(A160T), previously built. The expression of *yqhD* or *yqhD** (G149E) was realised by transforming the strains with a PCR fragment corresponding to a pTEF1-*yqhD*-CYCt-pTEF1-nat1-TEF1t cassette or pTEF1-*yqhD** (G149E)-CYCt-pTEF1-nat1-TEF1t cassette using the “short flanking homology” (SFH) method.
- 10 The pTEF1- *yqhD*-CYCt-pTEF1-nat1-TEF1t cassette or pTEF1-*yqhD** (G149E) -CYCt-pTEF1-nat1-TEF1t cassette were constructed using long PCR-based fusion of several fragments.
- pTEF1 and CYCt were amplified from the plasmid p405TEF1 (Addgene) using pTEF1/URA3 F and pTEF R primers and CYCt / *yqhD* F and CYCt / Nat1R primers respectively.
- 15 *yqhD* and *yqhD** were amplified respectively from pSCB-*yqhD* and pSCB *yqhD** (G149E) using primers *yqhD* / TEF-F and *yqhD* /CYCtR.
- pTEF1-nat1-TEF1t was amplified from the plasmid pAG35 from Euroscarf using Nat1/CYCt F and Nat1/Leu2 as primers.
- All fragments were amplified using primers having overlapping ends as described in Table
- 20 14. Each fragment was then purified.
- 100 ng of each fragment was used in a PCR experiment without primers using low annealing conditions allowing their simultaneous fusion.
- The unpurified product obtained in this step was used as a matrix in a PCR experiment with at high T_m , using pTEF1/ LEU2 F and Nat1/Leu2 primers having an extension of 40
- 25 bp homologous to the 40 first and 40 last bp of the *LEU2* locus (Table 14).
- These fragments were integrated in the *LEU2* locus, replacing the *LEU2* open reading frame.
- The transformation method used was the lithium acetate method. The strain CENPK, $\Delta gpd2$, *gldA**(A160T) was transformed either by pTEF1- *yqhD*-CYCt-pTEF1-nat1-TEF1t
- 30 cassette or by pTEF1-*yqhD** (G149E)-CYCt-pTEF1-nat1-TEF1t to obtained CENPK,

$\Delta gpd2$, *gldA** (A160T), *yqhD* and CENPK, $\Delta gpd2$, *gldA** (A160T), *yqhD** (G149E). Transformants were selected on YEPD rich medium (1% bacto yeast extract, 2% bacto peptone, 2% glucose) supplemented with 50 μ g/ml of nourseothricine (Weber bioagents, Germany). The integration of *yqhD* or *yqhD** (G149E) were confirmed by PCR on genomic DNA extracted, using YQHD ver F and YQHD ver R primers (Table 14). This resulted in the heterologous expression of *yqhD* and *yqhD** (G149E). The resulting strains were named CENPK $\Delta gpd2$, *gldA**(A160T), *yqhD* and CENPK $\Delta gpd2$, *gldA**(A160T), *yqhD**(G149E).

Table 14

Primer Name	Sequence	Description
pTEF1/ URA3 F	<u>ATG TCT GCC CCT AAG AAG ATC</u> <u>GTC GTT TTG CCA GGT GAC CAG</u> CTG GAG CTC ATA GCT TCA (SEQ ID NO 37)	The underlined 40 nucleotides correspond to the 40 first bp of <i>LEU2</i> gene 20 nucleotides in bold are homologous to the 20 first bp pTEF1
pTEF R	<u>TGC GGG TTG GGG TGT GCA GAT</u> <u>TAA AGT TGT TCA TAC TAG TGG</u> ATC CAC TAG TTC TAG AAA (SEQ ID NO 38)	The underlined 40 nucleotides correspond to the 40 first bp of <i>yqhD</i> or <i>yqhD*</i> (G419E) 20 nucleotides bold are homologous to the 20 last bp of pTEF1
<i>yqhD</i> / TEF-F	<u>CAT AGC AAT CTA ATC TAA GTT</u> <u>TTC TAG AAC TAG TGG ATC CAC</u> TAG TAT GAA CAA CTT TAA (SEQ ID NO 39)	The underlined 40 nucleotides correspond to the 40 last bp of pTEF1 20 nucleotides in bold are homologous to the 20 first bp of <i>yqhD</i> or <i>yqhD*</i> (G419E)
<i>yqhD</i> /CYCtR	<u>TGA AAT ATA AAT AAC GTT CTT</u> <u>AAT ACT AAC ATA ACT ATA AAC</u> TAG TTT AGC GGG CGG CTT (SEQ ID NO 40)	The underlined 40 nucleotides correspond to the 40 first bp of CYCt 20 nucleotides in bold are homologous to the 20 last bp of <i>yqhD</i> or <i>yqhD*</i> (G419E)
CYCt / <i>yqhD</i> F	<u>TGT CAG CCG CCG TAT ATA CGA</u> <u>AGC CGC CCG CTA AAC TAG TTT</u> ATA GTT ATG TTA GTA TTA (SEQ ID NO 41)	The underlined 40 nucleotides correspond to the 40 last bp of <i>yqhD</i> or <i>yqhD*</i> (G419E) 20 nucleotides in bold are homologous to the 20 first bp of CYCt
CYCt / Nat1R	<u>CTCCATGTCGCTGGCCGGGTGACCCGG</u> <u>CGGGGACGAGGCAGCAAATTAAA</u> GCCTTCGAGC (SEQ ID NO 42)	The underlined 40 nucleotides correspond to the 40 first bp of pTEF1 20 nucleotides in bold are homologous to the 20 last bp of CYCt
Nat1 /CYCt F	<u>GCT TGA GAA GGT TTT GGG ACG</u> <u>CTC GAA GGC TTT AAT TTG CTG</u> CCT CGT CCC CGC CGG GTC (SEQ ID NO 43)	The underlined 40 nucleotides correspond to the 40 last bp of CYCt 20 nucleotides in bold are homologous to the 20 first bp of pTEF1

Nat1 /Leu2	<u>TTA AGC AAG GAT TTT CTT AAC</u> <u>TTC TTC GGC GAC AGC ATC ACA</u> GTA TAG CGA CCA GCA TTC (SEQ ID NO 44)	The underlined 40 nucleotides correspond to the 40 last bp of <i>LEU2</i> gene 20 nucleotides in bold are homologous to the 20 last bp of TEF1t
YQHD ver F	ATG TCT GCC CCT AAG AAG ATC (SEQ ID NO 45)	20 nucleotides homologous to the 20 first bp of LEU2 gene
YQHD ver R	AC TAG TTT AGC GGG CGG CTT (SEQ ID NO 46)	20 nucleotides homologous to the 20 last bp of yqhD* (G149E)

1-3. Construction of two *S. cerevisiae* strain CENPK $\Delta gpd2$, $\Delta tpi1$, gldA*(A160T), yqhD, mgsA*(H21Q) and CENPK $\Delta gpd2$, $\Delta tpi1$, gldA*(A160T), yqhD*(G149E), mgsA*(H21Q)

5

The two strains used was CENPK, $\Delta gpd2$, gldA*(A160T), yqhD or CENPK, $\Delta gpd2$, gldA*(A160T), yqhD*, previously built.

The gene *TPII* was inactivated by transforming the strains with a PCR fragment corresponding to a pTEF1-hph-TEF1t-pPGK1-mgsA*(H21Q) cassette using the “short flanking homology” (SFH) method.

10 The pTEF1-hph-TEF1t-pPGK1-mgsA*(H21Q) cassette was constructed using long PCR-based fusion of several fragments.

The pTEF1-hph-TEF1t-pPGK1 were amplified from the plasmid pAG35pPGK1 constructed from pAG35 (Euroscarf) using PGK1/TPI1F and PGK1/ mgsAR

15 mgsA*(H21Q) was amplified from pETTOPO mgsA*(H21Q) using the primers mgsA/PGK1F and mgsA/TPI R as primers.

All fragments were amplified using primers having overlapping ends as described in Table 15. Each fragment was then purified.

100 ng of each fragment was used in a PCR experiment without primers and using low annealing conditions, allowing their simultaneous fusion.

20 The unpurified product obtained in this step was used as a matrix in a PCR experiment at high T_m , using PGK1/TPI1F and mgsA/TPI R primers having an extension of 40 bp homologous to the 40 first and 40 last bp of the *TPII* locus (Table 15).

This fragment was integrated in the *TPII* locus, replacing *TPII* open reading frame.

25 The transformation method used is the lithium acetate method. The strain CENPK, $\Delta gpd2$, gldA*(A160T), yqhD and the strain CENPK, $\Delta gpd2$, gldA*(A160T), yqhD* (G149E) was transformed by pTEF1-hph-TEF1t-pPGK1-mgsA*(H21Q) cassette to obtained CENPK, $\Delta gpd2$, $\Delta tpi1$, gldA*(A160T), yqhD* (G419E), mgsA*(H21Q), CENPK, $\Delta gpd2$, $\Delta tpi1$, gldA*(A16 Transformants were selected on YEPD rich medium (1% bacto yeast extract,

2% bactopectone, 2% glucose) supplemented with 250 µg/ml of hygromycin (Sigma-Aldrich).

The integration of *msgA**(H21Q) was confirmed by PCR on genomic DNA extracted, using *msgA* ver F and *msgA* ver R primers (Table 15).

- 5 This resulted in the heterologous expression of *msgA**(H21Q) and deletion of *TPI1*. The resulting strains were named CENPK $\Delta gpd2$, $\Delta tpi1$, *gldA**(A160T), *yqhD*, *msgA**(H21Q) and CENPK $\Delta gpd2$, $\Delta tpi1$, *gldA**(A160T), *yqhD**(G149E), *msgA**(H21Q).

Table 15

Primer Name	Sequence	Description
PGK1/TPI1F	<u>ATG TCG AAA GCT ACA TAT AAG</u> <u>GAA CGT GCT GCT ACT CAT CGC</u> CAG ATC TGT TTA GCT TGC (SEQ ID NO 47)	The underlined 40 nucleotides correspond to the 40 first bp of <i>TPI1</i> gene 20 nucleotides in bold are homologous to 20 the first bp pTEF1
PGK1/ <i>msgA</i> R	<u>GTG CGA GTC GTC AGT TCC ATA</u> ATA CGC AAA CCG CCT CTC C (SEQ ID NO 48)	The underlined 21 nucleotides correspond to the 40 first bp of <i>msgA</i> or <i>msgA</i> *(H21Q) 19 nucleotides in bold are homologous to 20 the last bp of pPGK1
<i>msgA</i> / PGK1F	<u>GGA GAG GCG GTT TGC GTA TTA</u> TGG AAC TGA CGA CTC GCA C (SEQ ID NO 49)	The underlined 22 nucleotides correspond to the 40 last bp of pPGK1 18 nucleotides in bold are homologous to the 20 first bp of <i>msgA</i> or <i>msgA</i> *(H21Q)
<i>msgA</i> /TPI R	<u>TTA GTT TTG CTG GCC GCA TCT</u> <u>TCT CAA ATA TGC TTC CCT TTA</u> CTT CAG ACG GTC CGC GAG (SEQ ID NO 50)	The underlined 40 nucleotides correspond to the 40 last bp of <i>TPI1</i> 20 nucleotides in bold are homologous to the 20 last bp of <i>msgA</i> or <i>msgA</i> *(H21Q)
<i>msgA</i> ver F	GG AAC TGA CGA CTC GCA C (SEQ ID NO 51)	20 nucleotides homologous to the 20 first bp of <i>TPI1</i> gene
<i>msgA</i> ver R	TTAGTTTCTAGAGTTGATGA (SEQ ID NO 52)	20 nucleotides homologous to the 20 last bp of <i>msgA</i> or <i>msgA</i> *(H21Q)

10

2 –Assesment of 1,2-propanediol production in *S. cerevisiae* CENPK $\Delta gpd2$, $\Delta tpi1$, *gldA**(A160T), *yqhD*, *msgA**(H21Q)

15

The strain CENPK, $\Delta gpd2$, $\Delta tpi1$, *gldA**(A160T), *yqhD*, *msgA**(H21Q) described above and the control strain CEN.PK2-1C were cultivated in batch culture, under anaerobic or aerobic conditions in minimal medium (SD medium 0.67% of Yeast nitrogen base without amino acid (DIFCO)) containing either 5% of glucose or 5% of sucrose as sole carbon source. Minimal medium was supplemented with 50 mg/l of uracil, 250 mg/l of leucine, 50

mg/L of histidine and 50 mg/L of tryptophan. Cultures were grown at 28°C under agitation at 225 rpm.

Aerobic cultures were carried out in shake flask of 250 ml containing 50 ml of medium. Anaerobic cultures were carried out in penicillin flask of 100 ml containing 90 ml of
5 medium.

At the end of the culture, 1,2-propanediol in the fermentation broth were analysed by gas chromatography/mass spectrometry (GC/MS) with an Agilent 7890A Series gas chromatograph coupled to an Agilent 5975C Series mass selective detector (EI) and a HP INNOWax column. The retention time and mass spectrum of 1,2-propanediol generated
10 were compared to those of authentic 1,2-propanediol. Residual glucose or sucrose in the fermentation broth were analysed by HPLC using a Biorad HPX 97H column for the separation and a refractometer for the detection. The yields of 1,2-propanediol over glucose or sucrose were then calculated.

Table 16: production of 1,2-propanediol in minimal medium in aerobic or anaerobic
15 conditions with glucose or sucrose as carbon source.

Strain	Conditions	Carbon source	1,2-propanediol titer (mg/l)	1,2-propanediol yield (mg/g carbon source)
CEN.PK2-1C	aerobic culture	glucose	44	0.9
CENPK, $\Delta gpd2$, $\Delta tpi1$, gldA*(A160T), yqhD, msgA*(H21Q)	aerobic culture	glucose	75	1.6
CEN.PK2-1C	aerobic culture	sucrose	41	0.9
CENPK, $\Delta gpd2$, $\Delta tpi1$, gldA*(A160T), yqhD, msgA*(H21Q)	aerobic culture	sucrose	69	1.6
CEN.PK2-1C	anaerobic culture	glucose	11	0.2
CENPK, $\Delta gpd2$, $\Delta tpi1$, gldA*(A160T), yqhD, msgA*(H21Q)	anaerobic culture	glucose	44	0.9

The production of 1,2-propanediol in a *S. cerevisiae* strain with a mutant MGS was improved under anaerobic or aerobic conditions with glucose or sucrose as compared with
20 the non-modified control strain.

REFERENCES

- Cooper RA and Anderson A (1970), *FEBS Lett.* **11** : 273-276
- 5 - Hopper DJ and Cooper RA (1971), *FEBS Lett.* **13** : 213-216
- Hopper DJ and Cooper RA (1972), *Biochem. J.* **128** : 321-329
- Cooper RA (1984), *Annu. Rev. Microbiol.* **38** : 49-68
- Saadat D and Harrison DHT (1998), *Biochemistry* **37** : 10074-10086
- Saadat D and Harrison DHT (1999), *Structure* **7** : 309-317
- 10 - Saadat D and Harrison DHT (2000), *Biochemistry* **39** : 2950-2960
- Marks GT, Susler M, Harrison DHT (2004), *Biochemistry* **43** : 3802-3813
- Töttemeyer S, Booth NA, Nichols WW, Dunbar B, Booth IR (1998), *Mol. Microbiol.* **27** : 553-562
- Ferguson GP, Töttemeyer S, MacLean MJ, Booth IR (1998), *Arch. Microbiol.* **170** : 209-219
- 15 - Garvie EI (1980), *Microbiol. Rev.* **44** : 106-139
- Cooper RA (1984), *Annu. Rev. Microbiol.* **38** : 49-68 *Arch. Microbiol.* **170**: 209-219
- Rule GS, Pratt EA, Chin CCQ, Wold F, Ho C (1985), *J. Bacteriol.* **161** : 1059-1068
- Dong JM, Taylor JS, Latour DJ, Iuchi S, Lin ECC (1993), *J. Bacteriol.* **175** : 6671-6678
- 20 - Grabar TB, Zhou S, Shanmugam KT, Yomano LP, Ingram LO (2006), *Biotechnol. Lett.* **28** : 1527-1535
- Misra K, Banerjee AR, Ray S, Ray M (1995), *Biochem. J.* **305** : 999-1003
- Cameron DC, Altaras NE, Hoffman ML, Shaw AJ (1998), *Biotechnol. Prog.* **14** : 116-125
- 25 - Altaras NE and Cameron DC (2000), *Biotechnol. Prog.* **16** : 940-946
- Bennett GN and San KY, (2001), *Appl. Microbiol. Biotechnol.* **55**: 1-9
- Ko J, Kim I, Yoo S, Min B, Kim K, Park C (2005), *J. Bacteriol.* **187** : 5782-5789
- Datta R and Henry M, (2006), *J. Chem. Technol. Biotechnol.* **81**: 1119-1129
- 30 - Wasewar KL, (2005), *Chem. Biochem. Eng. Q.* **19**: 159-172
- The UniProt consortium, (2008), *Nucleic Acids Res.* **36**: D190-195
- Guldener, U., et al., (1996), *Nucleic Acids Res.* **24**: 2519-24
- Schiestl, R. H. and Gietz, R. D., (1989), *Curr Genet.* **16**: 339-46.
- Shevchuk, N. A., et al., (2004), *Nucleic Acids Res.* **32**: e19

87100-42

52a

SEQUENCE LISTING IN ELECTRONIC FORM

In accordance with section 111(1) of the *Patent Rules*, this description contains a sequence listing in electronic form in ASCII text format (file: 87100-42 Seq 10-JAN-12 v1.txt).

A copy of the sequence listing in electronic form is available from the Canadian Intellectual Property Office.

The sequences in the sequence listing in electronic form are reproduced in the following table.

SEQUENCE TABLE

<110> METABOLIC EXPLORER
VOELKER, FRANCOIS
DUMON-SEIGNOVERT, LAURENCE
SOUCAILLE, PHILIPPE

<120> Mutant methylglyoxal synthase (MGS) for the production of a biochemical
by fermentation

<130> 87100-42

<140> PCT/EP2010/061094

<141> 2010-07-30

<150> EP09166815.2

<151> 2009-07-30

<150> US 61/230,076

<151> 2009-07-30

<160> 52

<170> PatentIn version 3.5

<210> 1

<211> 152

<212> PRT

<213> Escherichia coli

<400> 1

Met	Glu	Leu	Thr	Thr	Arg	Thr	Leu	Pro	Ala	Arg	Lys	His	Ile	Ala	Leu
1				5					10					15	

Val	Ala	His	Asp	His	Cys	Lys	Gln	Met	Leu	Met	Ser	Trp	Val	Glu	Arg
			20					25					30		

87100-42

52b

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

<210> 2
 <211> 152
 <212> PRT
 <213> Escherichia coli

<400> 2
 Met Glu Leu Thr Thr Arg Thr Leu Pro Ala Arg Lys His Ile Ala Leu
 1 5 10 15
 Val Ala His Asp Gln Cys Lys Gln Met Leu Met Ser Trp Val Glu Arg
 20 25 30
 His Gln Pro Leu Leu Glu Gln His Val Leu Tyr Ala Thr Gly Thr Thr
 35 40 45
 Gly Asn Leu Ile Ser Arg Ala Thr Gly Met Asn Val Asn Ala Met Leu
 50 55 60
 Ser Gly Pro Met Gly Gly Asp Gln Gln Val Gly Ala Leu Ile Ser Glu
 65 70 75 80
 Gly Lys Ile Asp Val Leu Ile Phe Phe Trp Asp Pro Leu Asn Ala Val
 85 90 95
 Pro His Asp Pro Asp Val Lys Ala Leu Leu Arg Leu Ala Thr Val Trp
 100 105 110
 Asn Ile Pro Val Ala Thr Asn Val Ala Thr Ala Asp Phe Ile Ile Gln
 115 120 125

87100-42

52c

Ser Pro His Phe Asn Asp Ala Val Asp Ile Leu Ile Pro Asp Tyr Gln
 130 135 140

Arg Tyr Leu Ala Asp Arg Leu Lys
 145 150

<210> 3
 <211> 152
 <212> PRT
 <213> Escherichia coli

<400> 3
 Met Glu Leu Thr Thr Arg Thr Leu Pro Ala Arg Lys His Ile Ala Leu
 1 5 10 15

Val Ala His Asp His Cys Lys Gln Met Leu Met Ser Trp Val Glu Arg
 20 25 30

His Gln Pro Leu Leu Glu Gln His Val Leu Tyr Ala Thr Gly Thr Thr
 35 40 45

Gly Asn Leu Ile Ser Arg Ala Thr Gly Met Asn Val Asn Ala Met Leu
 50 55 60

Ser Gly Pro Met Gly Gly Asp Gln Gln Val Gly Ala Leu Ile Ser Glu
 65 70 75 80

Gly Lys Ile Asp Val Leu Ile Phe Phe Trp Asp Pro Leu Asn Ala Val
 85 90 95

Pro His Asp Pro Asp Val Lys Ala Leu Leu Arg Leu Ala Thr Val Trp
 100 105 110

Asn Ile Pro Leu Ala Thr Asn Val Ala Thr Ala Asp Phe Ile Ile Gln
 115 120 125

Ser Pro His Phe Asn Asp Ala Val Asp Ile Leu Ile Pro Asp Tyr Gln
 130 135 140

Arg Tyr Leu Ala Asp Arg Leu Lys
 145 150

<210> 4
 <211> 23
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Oligonucleotide

<400> 4
 caccatggaa ctgacgactc gca

87100-42

52d

<210> 5
 <211> 23
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Oligonucleotide

<400> 5
 ttacttcaga cgggtccgcga gat

23

<210> 6
 <211> 48
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Oligonucleotide

<400> 6
 ttctgggatc cactaaatgt cgtgccgcac gatcctgacg tcaaagcc

48

<210> 7
 <211> 48
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Oligonucleotide

<400> 7
 ggctttgacg tcaggatcgt gcggcacgac atttagtgga tcccagaa

48

<210> 8
 <211> 42
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Oligonucleotide

<400> 8
 gcgacggtat ggaacattcc gctcgcgacc aacgtggcaa cg

42

<210> 9
 <211> 42
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Oligonucleotide

87100-42

52e

<400> 9
cggtgccacg ttggtcgcga gcggaatgtt ccataccgtc gc 42

<210> 10
<211> 45
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide

<400> 10
ctggtggcac acgatcaatg caaacagatg ctgatgagct gggtg 45

<210> 11
<211> 45
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide

<400> 11
caccagctc atcagcatct gtttgcattg atcgtgtgcc accag 45

<210> 12
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide

<400> 12
ccgacagtaa gacgggtaag cctg 24

<210> 13
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide

<400> 13
agcttagtaa agccctcgct ag 22

<210> 14
<211> 100
<212> DNA
<213> Artificial Sequence

87100-42

52f

<220>

<223> Oligonucleotide

<400> 14

catgggctag ctacgtatta attaaagatc tcctagggag ctcaccgggtt aaaaataaga	60
gttaccttaa atggtaactc ttattttttt aggcgcgcca	100

<210> 15

<211> 100

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide

<400> 15

agcttgggcgc gcctaaaaaa ataagagtta ccatttaagg taactcttat tttaaccgg	60
tgagctccct aggagatctt taattaatac gtagctagcc	100

<210> 16

<211> 43

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide

<400> 16

cgatgcacgt catgaacaac tttaatctgc acaccccaac ccg	43
---	----

<210> 17

<211> 29

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide

<400> 17

ctagctagct tagcgggagg cttcgtata	29
---------------------------------	----

<210> 18

<211> 48

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide

<400> 18

ggttcagaat ccaacgcaga agcggtgata tcccgtaaaa ccacaggc	48
--	----

87100-42

52g

<210> 19
 <211> 48
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Oligonucleotide

<400> 19
 gcctgtgggtt ttacgggata tcaccgcttc tgcgttggat tctgaacc 48

<210> 20
 <211> 45
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Oligonucleotide

<400> 20
 gacaccaaaa tcgtcgctgg cacacctgca cgtctgctag cggcg 45

<210> 21
 <211> 45
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Oligonucleotide

<400> 21
 cgccgctagc agacgtgcag gtgtgccagc gacgatTTTg gtgtc 45

<210> 22
 <211> 105
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Oligonucleotide

<400> 22
 gttaactacg gatgtacatt atggaactga cgactcgcac tttacctgcg cggaaacata 60
 ttgcgctggg ggcacacgat caggcctggg gatgatggcg ggatc 105

<210> 23
 <211> 120
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Oligonucleotide

87100-42

52h

<400> 23
 gggaaattaa gttaccggta gtgcctgttg catacagtac gtgttggttc agtaacggtt 60
 gatgccgttc caccagctc atcagcatct gtttgcattc agaagaactc gtcaagaagg 120

<210> 24
 <211> 100
 <212> DNA
 <213> Artificial sequence

<220>
 <223> Oligonucleotide

<400> 24
 tccagtcgcc gcatttcaac gacgcggtcg atattctgat ccccgattat cagcgttatc 60
 tcgcggaccg tctgaagtaa tgtaggctgg agctgcttcg 100

<210> 25
 <211> 100
 <212> DNA
 <213> Artificial sequence

<220>
 <223> Oligonucleotide

<400> 25
 tgtggaaata ctgaaaaatc tggatgtgcc ggtggcgaga aaaccgtaag aaacaggtgg 60
 cgtttgccac ctgtgcaata catatgaata tcctccttag 100

<210> 26
 <211> 41
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Ptrc01 promoter

<400> 26
 gagctgttga ctattaatca tccggctcga ataattgtgtg g 41

<210> 27
 <211> 61
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Oligonucleotide

<400> 27
 atgcttgctg tcagaagatt aacaagatac acattcctta gtttatcatt atcaatactc 60
 g 61

87100-42

52i

<210> 28
 <211> 23
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Oligonucleotide

<400> 28
 atcctcgaaa ctaagttctt ggt 23

<210> 29
 <211> 42
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Oligonucleotide

<400> 29
 aaacaccaga acttagtttc gaactagttt attcccactc tt 42

<210> 30
 <211> 62
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Oligonucleotide

<400> 30
 tgaaatataa ataacgttct taataactaac ataactataa actagtatgg accgcattat 60
 tc 62

<210> 31
 <211> 66
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Oligonucleotide

<400> 31
 cctggatgta ttaccggt gattgaataa tgcggtccat actagtttat agttatgtta 60
 gtatta 66

<210> 32
 <211> 60
 <212> DNA
 <213> Artificial Sequence

87100-42

52j

<220>

<223> Oligonucleotide

<400> 32

gaggcaagct aaacagatct ctagacctag gtacccgccg gcaaattaaa gccttcgagc 60

<210> 33

<211> 62

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide

<400> 33

gcttgagaag gttttgggac gctcgaaggc tttaatttgc taggtctaga gatctgttta 60
gc 62

<210> 34

<211> 64

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide

<400> 34

ctattcgtca tcgatgtcta gctcttcaat catctccggt ccactagtgg atctgatatc 60
acct 64

<210> 35

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide

<400> 35

atgcttgctg tcagaagatt 20

<210> 36

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide

<400> 36

tagtatggac cgcattattc 20

87100-42

52k

<210> 37
<211> 60
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide

<400> 37
atgtctgccc ctaagaagat cgtcgttttg ccaggtgacc agctggagct catagcttca 60

<210> 38
<211> 60
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide

<400> 38
tgcggttgg ggtgtgcaga ttaaagttgt tcatactagt ggatccacta gttctagaaa 60

<210> 39
<211> 60
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide

<400> 39
catagcaatc taatctaagt tttctagaac tagtggatcc actagtatga acaactttaa 60

<210> 40
<211> 60
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide

<400> 40
tgaaatataa ataacgttct taataactaac ataactataa actagtttag cgggcggcctt 60

<210> 41
<211> 60
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide

87100-42

521

<400> 41
 tgtcagccgc cgtatatacg aagccgcccgc ctaaactagt ttatagttat gttagtatta 60

<210> 42
 <211> 60
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Oligonucleotide

<400> 42
 ctccatgtcg ctggccgggt gacccggcgg ggacgaggca gcaaattaaa gccttcgagc 60

<210> 43
 <211> 60
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Oligonucleotide

<400> 43
 gcttgagaag gttttgggac gctcgaaggc ttttaatttgc tgccctcgtcc ccgccgggtc 60

<210> 44
 <211> 60
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Oligonucleotide

<400> 44
 ttaagcaagg attttcttaa cttcttcggc gacagcatca cagtatagcg accagcattc 60

<210> 45
 <211> 21
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Oligonucleotide

<400> 45
 atgtctgccc ctaagaagat c 21

<210> 46
 <211> 20
 <212> DNA
 <213> Artificial Sequence

87100-42

52m

<220>

<223> Oligonucleotide

<400> 46

actagttag cgggcggctt

20

<210> 47

<211> 60

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide

<400> 47

atgtcgaaag ctacatataa ggaacgtgct gctactcatc gccagatctg ttagcttgc

60

<210> 48

<211> 40

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide

<400> 48

gtgcgagtcg tcagttccat aatacgcaaa ccgcctctcc

40

<210> 49

<211> 40

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide

<400> 49

ggagaggcgg ttgctgtatt atggaactga cgactcgac

40

<210> 50

<211> 60

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide

<400> 50

ttagttttgc tggccgcatc ttctcaaata tgcttcctt tacttcagac ggtccgcgag

60

87100-42

52n

<210> 51
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide

<400> 51
ggaactgacg actcgcac

18

<210> 52
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide

<400> 52
ttagtttcta gagttgatga

20

CLAIMS

1. A mutant methylglyoxal synthase (MGS) enzyme, comprising an amino acid sequence selected from the group consisting of SEQ ID NO 1, SEQ ID NO 2 and SEQ ID NO 3.
2. A DNA molecule comprising a nucleotide sequence coding for the mutant MGS according to claim 1.
3. A modified microorganism, wherein said microorganism expresses the mutant methylglyoxal synthase (MGS) enzyme of claim 1.
4. The modified microorganism of claim 3, which is further modified for an improved production of a biochemical selected from lactic acid, acetol and 1,2-propanediol.
5. The modified microorganism of claim 4, for producing acetol, which additionally comprises a gene coding for a mutant YqhD, which has increased catalytic efficiency toward NADPH.
6. The modified microorganism of claim 4, for producing 1,2-propanediol, which additionally comprises a gene coding for a mutant YqhD which has increased catalytic efficiency toward NADPH, and/or a gene coding for a mutant glycerol dehydrogenase which has a reduced inhibition of its activity by NAD⁺ and/or its substrate and/or its product.
7. The modified microorganism of claim 5 or 6, wherein the mutant YqhD comprises at least one mutation selected from the group consisting of G149E, G149S, A286T and any combinations thereof, the amino acids position being given by reference to the YqhD sequence of *E.coli*.
8. The modified microorganism of claim 6, wherein the mutant glycerol dehydrogenase comprises at least one mutation selected from the group consisting of A160T, T120N and any combinations thereof, the amino acids position being given by reference to the GldA sequence of *E.coli*.

9. The modified microorganism of any one of claims 3 to 8, which is selected from the group consisting of *Enterobacteriaceae*, *Bacillaceae*, *Clostridiaceae*, *Streptomycetaceae*, *Corynebacteriaceae* and yeasts.

10. The modified microorganism of claim 9, which is selected from the group consisting of
5 *Escherichia coli*, *Klebsiella pneumoniae*, *Thermoanaerobacterium thermosaccharolyticum*, *Clostridium sphenoides* and *Saccharomyces cerevisiae*.

11. A method for the production of a biochemical selected from lactic acid, acetol and 1,2-propanediol, comprising:

- culturing a microorganism modified for an improved production of said
10 biochemical selected from the group consisting of lactic acid, acetol and 1,2-propanediol, in an appropriate culture medium, and
- recovery of the biochemical,

wherein the microorganism expresses a methylglyoxal synthase (MGS) enzyme which activity is not inhibited by orthophosphate, as defined in any one of claims 3 to 10.

12. The method of claim 11, wherein the microorganism is *Escherichia coli*.

13. The method of claims 11 or 12, wherein the recovered biochemical is further purified.

14. The method of any one of claims 11 to 13, wherein the appropriate culture medium comprises a carbon source selected from the group consisting of glucose, sucrose, mono- or disaccharides, starch and its derivatives, and any mixtures thereof.

15. The method of claim 14, wherein the source of carbon is selected from the group consisting of glucose and sucrose.

16. The method of any one of claims 11 to 15, wherein the production is done in a batch, fed-batch or continuous process.

1/4

1

MGSA_TRESO	-----MRR	KLRIALVAHD	NRKADIVDWA	LNNAEMLSQH	RLFGTGTGT	LVRESFMKRG	I-ASDITCMH
MGSA_COXBU	-----MT	VKKIALVAHD	RMKKELIEWI	KKHQNLLKHH	ELYATGSTGQ	AIEKT-----	L-NVTVTKME
MGSA_RALSO	-----MPK	RRRIALIAHD	HKKDDMIAFA	QTHKAPLMRC	DLATGTGTGG	RLQDE-----	V-GLSVQRMML
MGSA_BURMA	-----MS	TPRIALIAHD	AKKDDIVALA	GAYRATLAQC	RLVATGTGTGG	RIAQA-----	H-GLDVERKL
MGSA_BRUME	-----MTQ	RLRIALIAHD	QKKDDMVAF	RAHEQALSRY	DIVATGTGTGG	LIQDA-----	CPSLNIHRVK
MGSA_AGRT5	-----MEG	QRCIALIAHD	EKKDDMADFA	RHHQKVLASF	RIVATGTGTGG	RVQEA-----	CPGLEVIRLK
MGSA_RHIME	-----MAD	RKCLALIAHD	QKKDDLAAFA	KANEAVLSKW	KIVATGTGTGG	RVLDV-----	CPALDIVRLK
MGSA_SYMTH	-----	-MRIALIAHD	NRKQDMLKFV	KDHKSAREGH	QLFATGTGTGK	LIREH-----	T-GLDVHCFL
MGSA_BORBU	-----M	EKKIALIAHD	KKKEDLVNPF	KQNYLPLSKF	KLIATGTGTGS	KIQQA-----	T-DLTIFKYK
MGSA_BORGA	-----M	EKKIALIAHD	KKKDDLNVFV	KQNYLPLSKF	KLIATGTGTGS	KIQQA-----	T-DLTIIKYK
MGSA_CLOPE	-----	-MKIALIAHD	KKKEEMIELA	KDFEDKLSKH	ILVATGTGTGL	KIMQN-----	T-SLEVKRCK
MGSA_OCEIH	-----	-MNIALIAHD	EKKEDMIQFT	TAYTHVLSKH	RLFATGTGTGM	KISNA-----	T-GLQIHRFQ
MGSA_THETN	-----	-MNIALIAHD	QKKELMVNFA	IAYKHIFEKC	NLYATGHTGQ	LIKEA-----	T-GLDVHCCLL
MGSA_CLOTS	-----MV	NLNIALIAHD	MKKSLMIDFA	IAYKDILEKC	NIYATGATGQ	LVEEA-----	T-GIKVKNFL
MGSA_GEOKA	-----	-MKIALIAHD	QKKEEMVAF	TAYAPVLANH	ELYATGTGTGL	RIQEA-----	T-GLPVHRFQ
MGSA_BACCR	-----	-MKIALIAHD	KKKNDMVSFA	YAYKPIFEQH	ELFATGTGTGL	RIMEA-----	T-GLVITRYQ
MGSA_BACSK	-----	-MKIALIAHD	KKKQELVDFC	VAYEPILKEH	ELYATGTGTGT	RIMEA-----	T-ELVVTRFK
MGSA_BACHD	-----	-MRIALIAHD	KKKDEMVOFT	IAYKDVLDKH	ELYATGTGTGM	RIMEA-----	A-QLPVKRFK
MGSA_BACSU	-----	-MKIALIAHD	KKKQDMVOFT	TAYRDILKNH	DLYATGTGTGL	KIHEA-----	T-GLQIERFQ
MGSA_ENTFA	-----	-MKIALIAHD	RKKTLMIKLA	TAYKHILEKH	ELYATGTGTGM	KVMEA-----	T-GLPVHCFK
MGSA_LISIN	-----	-MHIALIAHD	EKKDLMVGFA	TAYKHILLEPH	QLYATGTGTGL	RIIEA-----	T-GLTVHRFK
MGSA_HAEIN	MQTTTRTLTQ	HKRIALVAHD	SCKKNLLNWT	QKHKEALKPH	ILYATGTGTGH	ILERE-----	T-GLSIQSLL
MGSA_PHOPR	MQVTTRTMNK	SKHIALVAHD	NCKQDLLRWV	KEFEDKLEDH	TLYATGTGTGH	LLSKE-----	T-GLKINCMI
MGSA_VIBVU	MQKTTRTMPA	HKHIALVAHD	NYKPELLRWV	KENKEKLQSH	FLYATGTGTGH	MLSRE-----	T-GLAIKSMI
MGSA_VIBCH	MKKTTRTMAA	HKHVALVAHD	NCKGELLRWV	TENKEKLQRH	FLYATGTGTGH	MLSKE-----	T-GLAIKSMI
MGSA_VIBPA	MQKTTRTMPA	HKHVALVAHD	NCKPELLRWV	KENKEKLQRH	FLYATGTGTGH	MLSKE-----	T-GLAIKSMI
MGSA_PHOLL	MELTTRTMAT	AKNIALIAHD	HCKTSLLAWS	KKHKSLLKKH	HLYATGTGTGN	LIQNE-----	T-GLSVTNML
MGSA_YERPE	MELTTRTIAA	RKHIALVSHD	HCKKSLLAWS	MENRDLLAQH	ELYATGTGTGN	LVQKA-----	T-GIDVHCCLL
MGSA_ERWCT	MEFTTRTIPA	QKHIALVAHD	HCKQSLLDWV	GTNKQQLTEH	TLYATGTGTGN	LIQSN-----	T-GLPVKSML
MGSA_ECOL6	MELTTRTLPS	RKHIALVAHD	HCKQMLMSWV	ERHQPLLEQH	VLYATGTGTGN	LISRA-----	T-GMNVNAML
MGSA_SALTI	MELTTRTLPT	RKHIALVAHD	HCKQMLMNWV	ERHQPLLEKH	VLYATGTGTGN	LIQRA-----	T-GMDVNAML
MGSA_MANSN	METTFRHVAA	QKHIALVAHD	HCKEDLINWC	QKNVHHLQNH	QLYATGTGTGH	LIEKA-----	T-ELKINSLL
MGSA_PASMU	MQSTARTLSV	NKHIALVAHD	HCKQDLINWC	KTHRTLLAQH	TLYATGTGTGN	LIQKE-----	A-NLPIKSLL
MGSA_CLOAB	-----MNS	KKKIALVAHD	NRKKALISWC	EANSEVLSNH	SLCGTGTAK	LIKEA-----	T-GLVFPYK
MGSA_THEMA	-----MDK	KKRIALIAHD	RRKRDLEWV	SFNLGTLSKH	ELYATGTGTGA	LLQEK-----	L-GLKVHRLK
MGSA_LEPIC	--MKEVSVPA	IKRIVLIAHD	NRKEDLVNWW	KTHREILSKH	QLYGTGTGTGK	LISEE-----	T-ELPVYRFL

71

MGSA_TRESO	SGPMGGDAEI	AALVVRKEID	FAVFFIDDLN	PQPHEADIQM	LLRQCRIHNI	PIACNRYSD	LMITSSLWDD
MGSA_COXBU	SGPLGGDLQL	GAKIVNKEID	ILIFFWDPLE	AQPHDPDVRA	LLRIAVVWNL	PVACNASTAD	YLLT-SPLFD
MGSA_RALSO	SGPWGGDLQI	GAQLAEGRVD	AVIFLRDPMT	PQPHEPDINA	LVRACDVHNI	PCATNLATAD	LVM-----
MGSA_BURMA	SGPLGGDLQI	GAELADGRVD	IVIFLRDPMT	AQPHDPDITA	LVRACDVHDV	PVATNVATAR	VLL-----
MGSA_BRUME	SGPLGGDQOI	GAMIAECTVE	VLIFFIDPLS	PLPHDVDVKA	LTRLGVSVDI	PMALNRATAE	KLV-----
MGSA_AGRT5	SGPLGGDQOI	GAMIATGEVD	MLIFFTDPLT	AMPHDVDVKA	LTRLATVYDI	PMALNRATAE	NLI-----
MGSA_RHIME	SGPLGGDQOI	GALIAATGVD	CLIFFVDPLT	AMPHDVDVKA	LMRLAIVYDI	PMALNRATAE	QLI-----
MGSA_SYMTH	SGPLGGDQOI	GSRVATGEID	MVIFLRDPLT	AMPHEPDVQ	LLRLCDVRDI	PVATNLGSAR	MFA-----
MGSA_BORBU	SGPMGGDQOI	GAEVAEGNVL	AIFFFRDPLT	SQPHEPDVSA	LIRLCDVHKI	PLATNVKTAE	ILI-----
MGSA_BORGA	SGPMGGDQOI	GAEVAEGNVL	AIFFFRDPLT	NQPHEPDVSA	LIRLCDVHNI	PLATNVKTAE	ILI-----
MGSA_CLOPE	SGPLGGDQEI	GAMVANHDVD	MVIFLRDPLT	AQPHEPDISA	LLRLCDVYKV	PLATNTESAK	LIM-----
MGSA_OCEIH	SGPLGGDQOI	GAMIANGEMD	MIIFFRDPLT	AQPHEPDVSA	LMRLCDVHQI	PLVTNIAGAE	IFI-----
MGSA_THETN	PGPLGGDQOI	GALIAENKID	LVIFLRDPLT	VQPHEPDILA	LLRVCDVHSI	PVATNIATAE	VLL-----
MGSA_CLOTS	PGPMGGDQOI	GAMIAEDKMD	LVIFLRDPLT	AQPHEPDILA	LLRVCDVHSI	PLATNLATAE	VLI-----
MGSA_GEOKA	SGPYGGDQEI	GAMIARNEMD	LVIFFRDPLT	AQPHEPDISA	LMRLCDVYAV	PLATNIGTAE	LLI-----
MGSA_BACCR	SGPLGGDQEI	GAMIAKNLDL	MVIFFRDPLT	AQPHEPDVNA	LLRLCDVYAI	PLATNMASAE	MLM-----
MGSA_BACSK	SGPLGGDQOI	GALVAENQFD	LILFMRDPLT	AQPHEPDVTA	LIRLCDVQSV	PLATNMGTAE	ILI-----
MGSA_BACHD	SGPLGGDQOI	GALVAENKFD	LIIFLRDPLT	AQPHEPDVTA	LVRLCDVQSV	PLATNIGTAE	ILI-----
MGSA_BACSU	SGPLGGDQOI	GALIAANALD	LVIFLRDPLT	AQPHEPDVSA	LIRLCDVYSI	PLATNMGTAE	ILV-----
MGSA_ENTFA	SGPLGGDQOI	GAMISEDNID	LVIFLRDPLS	AQPHEPDVTA	LIRLSDVYEI	PLATNIGSAE	ILL-----
MGSA_LISIN	SGPLGGDQOI	GARISENKMD	LVIFLRDPLT	AQPHEPDVSA	LIRLCDVYEI	PLATNIGTAE	ILI-----
MGSA_HAEIN	SGPMGGDQQL	GGLIAEKKID	MMIFFWXPMM	AAPHEPDVKA	LMRIATVWNI	PVAINQSSAD	FLLT-SVLFE
MGSA_PHOPR	SGPMGGDQQL	GALISESKID	MMIFFWDPLN	AVPHDPDVKA	LLRISAVWNV	PVATNRASAK	FMIT-SPLLA
MGSA_VIBVU	SGPMGGDQQL	GALISEGKID	MLIFFWDPLN	AVPHDPDVKA	LLRIASVWNI	PVATNRATAK	FLFE-SPLLN
MGSA_VIBCH	SGPMGGDQQL	GALISEGKID	VLIFFWDPLN	AVPHDPDVKA	LLRIASVWNI	PVATNRATAK	FLFS-SSLME
MGSA_VIBPA	SGPMGGDQQL	GALISEGKID	VLVFFWDPLN	AVPHDPDVKA	LLRIASVWNI	PVATNRATAK	FLFD-SPLLE
MGSA_PHOLL	SGPMGGDQOI	GSLISEGKID	VLIFFWDPLN	SVPHDPDVKA	LLRLATVWNI	PVATNLASAD	FIVS-SPLFS
MGSA_YERPE	SGPMGGDQEV	GALISEKKID	ILIFFWDPLN	AVPHDPDVKA	LLRLATVWNI	PVATNRSTAD	FLIG-STLFS
MGSA_ERWCT	SGPMGGDQOV	GALISEGKID	LMIFFWDPLN	AVPHDPDVKA	LLRLATVWNI	PVATNRATAD	FLIN-SALFK
MGSA_ECOL6	SGPMGGDQOV	GALISEGKID	VLIFFWDPLN	AVPHDPDVKA	LLRLATVWNI	PVATNVATAD	FIIQ-SPHFN
MGSA_SALTI	SGPMGGDQOV	GALISEGKID	VLIFFWDPLN	AVPHDPDVKA	LLRLATVWNI	PVATNVSTAD	FIIQ-SPHFN
MGSA_MANSN	SGPMGGDQQL	GALIAENKID	VMIFFWDPMN	AVPHDPDVKA	LLRIAASVWNI	PHAMNIASAD	LLIN-SPLIN
MGSA_PASMU	SGPMGGDQQL	GGLIAEKQID	VLIFFWDPMN	AVPHDPDVKA	LMRIATVWNI	PVAMNMATAD	FLIT-SPSFA
MGSA_CLOAB	SGPMGGDQOI	GAAIVNEDID	FMIFFWDPLT	AQPHDPDVKA	LLRISVLYDI	PIAMNESTAE	FLIK-SPIMK
MGSA_THEMA	SGPLGGDQOI	GAMIAEKKID	VLIFFWDPLE	PQAHDPDVKA	LIRIATVYNI	PVAITRSTAD	FLIS-SPLMN
MGSA_LEPIC	SGPLGGDQOI	GAKIAEGDLD	IVIFFWDPLT	AQPHDPDVKA	LLRIAVLYNV	PMACNRSTAD	YMIS-SPQFT

Fig. 1

2/4

141

MGSA_TRESO	AGYVPKDPIY	APFDRKAFEE	SLKVKE---
MGSA_COXBU	SDYHPETPDY	EAYRNRII--	-----
MGSA_RALSO	IALGLAQDPD	KEIHA-----	-----
MGSA_BURMA	DDLARRLTAN	A-----	-----
MGSA_BRUME	RALD-----	-----	-----
MGSA_AGRT5	-----DF	NSAD-----	-----
MGSA_RHIME	-----DF	RRN-----	-----
MGSA_SYMTH	DDL----MRL	KDVK-----	-----
MGSA_BORBU	KGLESLIF--	-----	-----
MGSA_BORGA	KGFEGLNT--	-----	-----
MGSA_CLOPE	ADI-----	-----	-----
MGSA_OCEIH	HGLSRGDLKW	REIIKERQEK	EGTP-----
MGSA_THETN	KGMEMGLLDW	RQI-----	-----
MGSA_CLOTS	KGLDAGLLEW	RNAV-----	-----
MGSA_GEOKA	RALERGDLAW	RSIVRGQTKG	GEESKTER-
MGSA_BACCR	HALERGDLDY	RKLR-----	-----
MGSA_BACSK	KGLERGDFAF	RDIIHEQEKA	NPLKEG---
MGSA_BACHD	KGLARGDFTW	RELVHDEEEP	RV-----
MGSA_BACSU	RTLDEGVFEF	RDLLRGEEP	NV-----
MGSA_ENTFA	RGVEAGFADF	REVIHEGDRR	PLAF-----
MGSA_LISIN	RGLGAGFLDW	RDLRRNDE--	-----
MGSA_HAEIN	QDVEIDVPDY	EGYLKERLA-	-----
MGSA_PHOPR	EEISIEIPDY	EAYLAERIG-	-----
MGSA_VIBVU	EEVDVEIPDY	QAYLAGRT--	-----
MGSA_VIBCH	QEVQIEIPDY	QAYLAERT--	-----
MGSA_VIBPA	QEV DIEVPDY	EAYLAERM--	-----
MGSA_PHOLL	ESVDIQVPDY	QHYLNGRLK-	-----
MGSA_YERPE	SEVTIAIPDY	DRYMQQRLDL	K-----
MGSA_ERWCT	EPVQIAIPDY	QRYLQDRLK-	-----
MGSA_ECOL6	DAVDILIPDY	QRYLADRLK-	-----
MGSA_SALTI	DAVDILIPDY	ARYLAERLK-	-----
MGSA_MANSM	REIELRIPDY	QTYLQKRLK-	-----
MGSA_PASMU	QETQVRIPDY	DGYLKERLK-	-----
MGSA_CLOAB	EQHERHIIDY	YQKIRKDNF-	-----
MGSA_THEMA	DVYEKIQIDY	EEELERRIRK	VVEGEEEEET
MGSA_LEPIC	KTYKKILLSY	NTKVKKD---	-----

Fig. 1 (end)

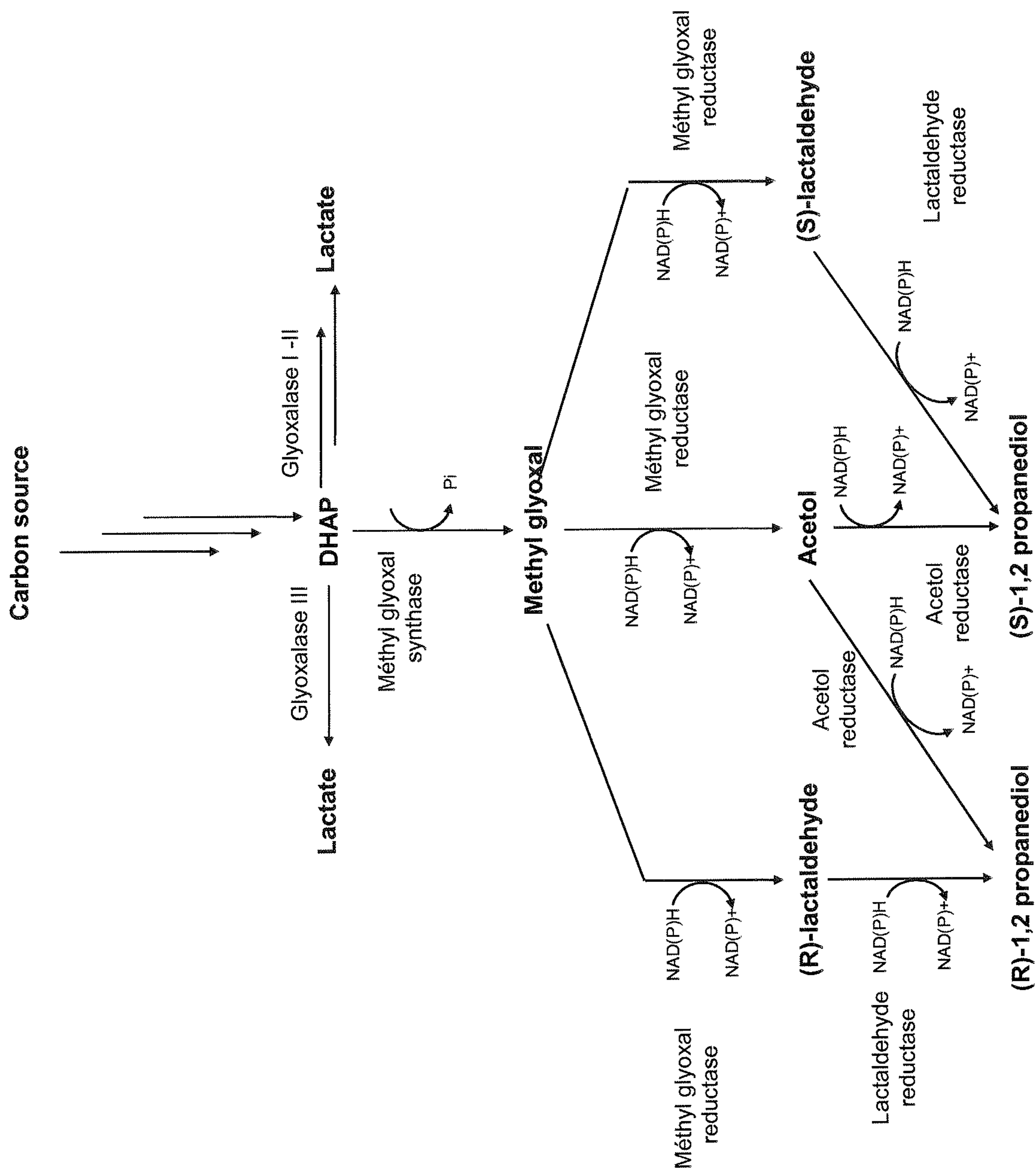


Fig. 2

4/4

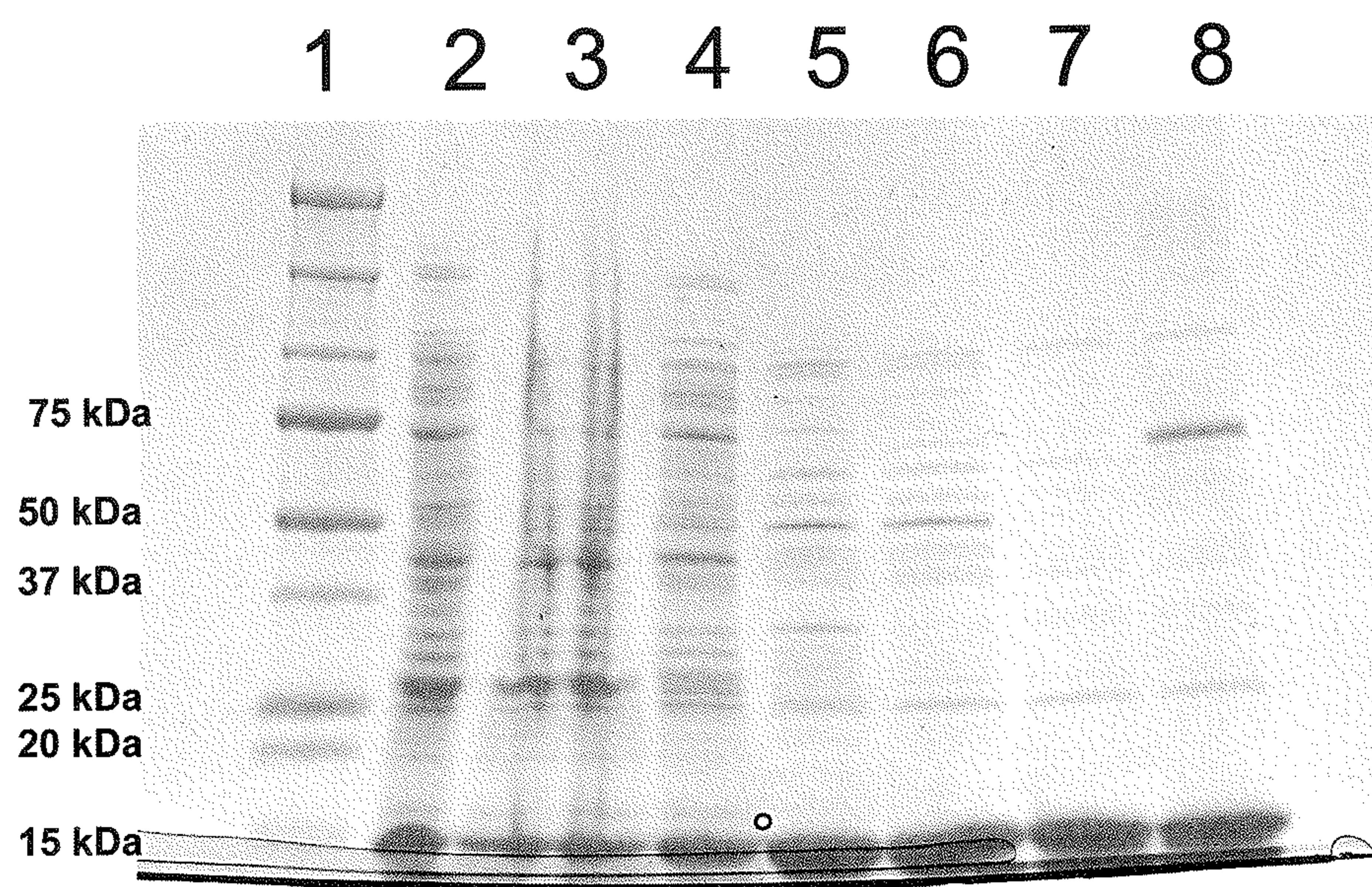


Fig. 3