TITLE

VARIANT SUCROSE TRANSPORTER POLYPEPTIDES

FIELD OF THE INVENTION

The invention relates to the fields of microbiology and molecular biology. More specifically, variant sucrose transporter polypeptides that enable bacterial growth over a wide range of gene expression levels and sucrose concentrations, recombinant bacteria comprising these variant sucrose transporter polypeptides, and methods of utilizing such bacteria to produce products such as glycerol and glycerol-derived products are provided.

BACKGROUND OF THE INVENTION

Many commercially useful microorganisms use glucose as their main carbohydrate source. However, a disadvantage of the use of glucose by microorganisms developed for production of commercially desirable products is the high cost of glucose. The use of sucrose and mixed feedstocks containing sucrose and other sugars as carbohydrate sources for microbial production systems would be more commercially desirable because these materials are usually readily available at a lower cost.

A production microorganism can function more efficiently when it can utilize any sucrose present in a mixed feedstock. Therefore, when a production microorganism does not have the ability to utilize sucrose efficiently as a major carbon source, it cannot operate as efficiently. For example, bacterial cells typically show preferential sugar use, with glucose being the most preferred. In artificial media containing mixtures of sugars, glucose is typically metabolized to its entirety ahead of other sugars. Moreover, many bacteria lack the ability to utilize sucrose. For example, less than 50% of *Escherichia coli* (*E. coli*) strains have the ability to utilize sucrose. Thus, when a production microorganism cannot

utilize sucrose as a carbohydrate source, it is desirable to engineer the microorganism so that it can utilize sucrose.

Recombinant bacteria that have been engineered to utilize sucrose by incorporation of sucrose utilization genes have been reported. For example, Livshits et al. (U.S. Patent No. 6,960,455) describe the production of amino acids using *Escherichia coli* strains containing genes encoding a metabolic pathway for sucrose utilization. Additionally, Olson et al. (*Appl. Microbiol. Biotechnol.* 74:1031-1040, 2007) describe *Escherichia coli* strains carrying genes responsible for sucrose degradation, which produce L-tyrosine or L-phenylalanine using sucrose as a carbon source. Eliot et al. (U.S. Patent Application No. 2011/0136190 A1) describe recombinant bacteria that produce glycerol and glycerol-derived products from sucrose.

However, problems remain in engineering production microorganisms so that they can utilize sucrose effectively. Specifically, high levels of expression of sucrose transport genes result in poor growth on sucrose because excess sucrose transport is inhibitory. On the other hand, low levels of sucrose transport also result in sub-optimal growth on sucrose. Therefore, it is difficult to obtain the proper sucrose transporter gene expression level. Additionally, expression of sucrose transport genes under conditions at which sucrose transport is in excess, such as at high sucrose concentrations, may inhibit growth even at gene expression levels at which growth is not inhibited at lower sucrose concentrations. Therefore, a need also exists for a sucrose transporter that can enable growth on sucrose over a broad range of sucrose concentrations.

SUMMARY OF THE INVENTION

One aspect of the disclosure provides a variant sucrose transporter polypeptide having an amino acid sequence that has at least 95% identity to an amino acid sequence as set forth in SEQ ID NO:26 based on

a Clustal W method of alignment and having an amino acid change from arginine to alanine or arginine to leucine at position 300, and comprising:

(a) at least one additional amino acid change selected from the group consisting of:

(i) glutamine to histidine at position 353

(ii) leucine to proline at position 61;

(iii) phenylalanine to leucine at position 159;

(iv) glycine to cysteine at position 162;

(v) proline to histidine at position 169;

(vi) leucine to tryptophan at position 61;

(vii) leucine to histidine at position 61;

(viii) leucine to phenylalanine at position 61; and

(ix) leucine to tyrosine at position 61; or

(b) a length of 402 to 407 amino acids from the N-terminus; or

(c) a length of 402 to 407 amino acids from the N-terminus, and having at least one of the amino acid changes of (a).

Another aspect provides a variant sucrose transporter polypeptide having an amino acid sequence that has at least 95% identity based on a Clustal W method of alignment to an amino acid sequence selected from the group consisting of SEQ ID NOs: 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, and 98, and comprising an amino acid at an equivalent position when compared with a reference amino acid sequence of SEQ ID NO:26 selected from the group consisting of:

(a) alanine at a position equivalent to position 300; and

(b) leucine at a position equivalent to position 300.

Another aspect provides a recombinant bacterium comprising in its genome or on at least one recombinant construct:

(a) a nucleotide sequence encoding a variant sucrose transporter polypeptide having an amino acid sequence that has at least 95% identity based on a Clustal W method of alignment to an amino acid sequence selected from the group consisting of SEQ ID NOs:26, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, and 98,and

an amino acid at an equivalent position when compared with a reference amino acid sequence of SEQ ID NO:26 selected from the group consisting of:

(i) alanine at a position equivalent to position 300; and

(ii) leucine at a position equivalent to position 300; and

(b) a nucleotide sequence encoding a polypeptide having sucrose hydrolase activity;

wherein (a) and (b) are each operably linked to the same or a different promoter, further wherein said recombinant bacterium is capable of metabolizing sucrose.

In one embodiment, the recombinant bacterium produces 1,3-propanediol, glycerol, and/or 3-hydroxypropionic acid.

Another aspect provides a process for making glycerol, 1,3-propanediol and/or 3-hydroxypropionic acid from sucrose comprising:

a) culturing the recombinant bacterium that produces 1,3-propanediol, glycerol, and/or 3-hydroxypropionic acid, disclosed herein, in the presence of sucrose; and

b) optionally, recovering the glycerol, 1,3-propanediol and/or 3-hydroxypropionic acid produced.

BRIEF SEQUENCE DESCRIPTIONS

The following sequences conform with 37 C.F.R. 1.821 1.825 ("Requirements for Patent Applications Containing Nucleotide Sequences and/or Amino Acid Sequence Disclosures - the Sequence Rules") and consistent with World Intellectual Property Organization (WIPO) Standard ST.25 (2009) and the sequence listing requirements of the EPO and PCT (Rules 5.2 and 49.5(a bis), and Section 208 and Annex C of the Administrative Instructions). The symbols and format used for nucleotide and amino acid sequence data comply with the rules set forth in 37 C.F.R. §1.822.

Table 1

Summary of Gene and Protein SEQ ID Numbers

Gene	Coding Sequence SEQ ID NO:	Encoded Protein SEQ ID NO:
GPD1 from *Saccharomyces cerevisiae*	1	2
GPD2 from *Saccharomyces cerevisiae*	3	4
GPP1 from *Saccharomyces cerevisiae*	5	6
GPP2 from *Saccharomyces cerevisiae*	7	8
dhaB1 from *Klebsiella pneumoniae*	9	10
dhaB2 from *Klebsiella pneumoniae*	11	12
dhaB3 from *Klebsiella pneumoniae*	13	14
aldB from *Escherichia coli*	15	16
aldA from *Escherichia coli*	17	18
aldH from *Escherichia coli*	19	20
galP from *Escherichia coli*	21	22
cscB from *Escherichia coli* EC3132	23	24
cscB from *Escherichia coli* ATCC®13281	25	26
cscA from *Escherichia coli* EC3132	27	28
cscA from *Escherichia coli* ATCC13281	29	30
bfrA from *Bifidobacterium lactis* strain DSM 10140^T	31	32
SUC2 from *Saccharomyces cerevisiae*	33	34
scrB from *Corynebacterium glutamicum*	35	36
sucrose phosphorylase gene from *Leuconostoc mesenteroides* DSM 20193	37	38
sucP Bifidobacterium adolescentis DSM 20083	39	40
scrK from *Agrobacterium tumefaciens*	41	42
scrK from *Streptococcus mutans*	43	44
scrK From *Escherichia coli*	45	46
scrK from *Klebsiella pneumoniae*	47	48
cscK from *Escherichia coli*	49	50
cscK from *Enterococcus faecalis*	51	52
HXK1 from *Saccharomyces cerevisiae*	53	54
HXK2 from *Saccharomyces cerevisiae*	55	56
dhaT from *Klebsiella pneumoniae*	57	58
dhaX from *Klebsiella pneumoniae*	59	60
scrT1 from *Citrobacter sp*	67	68
scrT3 from *Enterococcus faecium*	69	70
scrT4 from *Corynebacterium glucuronolyticum*	71	72
scrT5 from *Bifidobacterium animalis subsp. lactis*	73	74
scrT6 from *Bifidobacterium gallicum*	75	76
scrT7 from *Bifidobacterium longum*	77	78
scrT8 from *Bifidobacterium adolescentis*	79	80
scrT9 from *Bifidobacterium longum*	81	82

scrT12 from *Mitsuokella multacida*	83	84
scrT13 from *Lactobacillus antri*	85	86
scrT14 from *Lactobacillus ruminis*	87	88
scrT21 from *Yersinia frederiksenii*	89	90
scrT25 from *Serratia proteamaculans*	91	92
scrT26 from *Escherichia coli*	93	94
fruP from *Bacillus licheniformis* 14580	95	96
lacY from *Pseudomonas fluorescens* Pf5	97	98
cscB from *Escherichia coli* ATCC®13281 with R300A mutation	99	100
cscB from *Escherichia coli* ATCC®13281 with R300L mutation	101	102
cscB from *Escherichia coli* ATCC®13281 with R300A and Q353H mutations	103	104
cscB from *Escherichia coli* ATCC®13281 with R300A, Q353H, L61P mutations	105	106
scrT1 from *Citrobacter sp* with R305A mutation	107	108
scrT1 from *Citrobacter sp* with R305L mutation	109	110
scrT7 from *Bifidobacterium longum* with R312A mutation	111	112
scrB from *Pseudomonas fluorescens* Pf5	133	134
fruA from *Bacillus licheniformis* 14580	135	136

SEQ ID NO:61 is the nucleotide sequence of the cscAKB gene cluster from *Escherichia coli* ATCC®13281.

SEQ ID NO:62 is the nucleotide sequence of plasmid pSYCO101.

SEQ ID NO:63 is the nucleotide sequence of plasmid pSYCO103.

SEQ ID NO:64 is the nucleotide sequence of plasmid pSYCO106.

SEQ ID NO:65 is the nucleotide sequence of plasmid pSYCO109.

SEQ ID NO:66 is the nucleotide sequence of plasmid pSYCO400/AGRO.

SEQ ID NO:113 is the nucleotide sequence of plasmid pDMWP1.

SEQ ID NO:114 is the nucleotide sequence of plasmid pDMWP3.

SEQ ID NO:119 is the nucleotide sequence of plasmid pBHR-cscBKA.

SEQ ID NO:124 is the nucleotide sequence of the promoter/MCS/double terminator insert described in Examples 22-24.

SEQ ID NO:125 is the codon optimized nucleotide sequence of the coding region of *scrT1* for expression in *E. coli*.

SEQ ID NO:130 is the codon optimized nucleotide sequence of the coding region of *scrT7* for expression in *E. coli*.

SEQ ID NOs:115-118, 120-123, 126-129 and 131-132 are the nucleotide sequences of primers used in the Examples herein.

DETAILED DESCRIPTION

The disclosure of each reference set forth herein is hereby incorporated by reference in its entirety.

As used herein and in the appended claims, the singular forms "a", "an", and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, reference to "a cell" includes one or more cells and equivalents thereof known to those skilled in the art, and so forth.

In the context of this disclosure, a number of terms and abbreviations are used. The following definitions are provided.

"Open reading frame" is abbreviated as "ORF".

"Polymerase chain reaction" is abbreviated as "PCR".

"American Type Culture Collection" is abbreviated as "ATCC".

The term "recombinant glycerol-producing bacterium" refers to a bacterium that has been genetically engineered to be capable of producing glycerol and/or glycerol-derived products.

The term "polypeptide having sucrose transporter activity" refers to a polypeptide that is capable of mediating the transport of sucrose into microbial cells.

The term "polypeptide having fructokinase activity" refers to a polypeptide that has the ability to catalyze the conversion of D-fructose + ATP to fructose-phosphate + ADP. Typical of fructokinase is EC 2.7.1.4. Enzymes that have some ability to phosphorylate fructose, whether or not this activity is their predominant activity, may be referred to as a fructokinase. Abbreviations used for genes encoding fructokinases and proteins having fructokinase activity include, for example, "*Frk*", "*scrK*", "*cscK*", "*FK*", and "*KHK*". Fructokinase is

encoded by the *scrK* gene in *Agrobacterium tumefaciens* and *Streptococcus mutans*; and by the *cscK* gene in certain *Escherichia coli* strains.

The term "polypeptide having sucrose hydrolase activity" refers to a polypeptide that has the ability to catalyze the hydrolysis of sucrose to produce glucose and fructose. Such polypeptides are often referred to as "invertases" or "β-fructofuranosidases".

The terms "glycerol derivative" and "glycerol-derived products" are used interchangeably herein and refer to a compound that is synthesized from glycerol or in a pathway that includes glycerol. Examples of such products include 3-hydroxypropionic acid, methylglyoxal, 1,2-propanediol, and 1,3-propanediol.

The term "microbial product" refers to a product that is microbially produced, i.e., the result of a microorganism metabolizing a substance. The product may be naturally produced by the microorganism, or the microorganism may be genetically engineered to produce the product.

The terms "phosphoenolpyruvate-sugar phosphotransferase system", "PTS system", and "PTS" are used interchangeably herein and refer to the phosphoenolpyruvate-dependent sugar uptake system.

The terms "phosphocarrier protein HPr" and "PtsH" refer to the phosphocarrier protein encoded by *ptsH* in *E. coli*. The terms "phosphoenolpyruvate-protein phosphotransferase" and "PtsI" refer to the phosphotransferase, EC 2.7.3.9, encoded by *ptsI* in *E. coli*. The terms "glucose-specific IIA component", and "Crr" refer to enzymes designated as EC 2.7.1.69, encoded by *crr* in *E. coli*. PtsH, PtsI, and Crr comprise the PTS system.

The term "PTS minus" refers to a microorganism that does not contain a PTS system in its native state or a microorganism in which the PTS system has been inactivated through the inactivation of a PTS gene.

The terms "glycerol-3-phosphate dehydrogenase" and "G3PDH" refer to a polypeptide responsible for an enzyme activity that catalyzes the conversion of dihydroxyacetone phosphate (DHAP) to glycerol 3-

phosphate (G3P). *In vivo* G3PDH may be NAD- or NADP-dependent. When specifically referring to a cofactor specific glycerol-3-phosphate dehydrogenase, the terms "NAD-dependent glycerol-3-phosphate dehydrogenase" and "NADP-dependent glycerol-3-phosphate dehydrogenase" will be used. As it is generally the case that NAD-dependent and NADP-dependent glycerol-3-phosphate dehydrogenases are able to use NAD and NADP interchangeably (for example by the enzyme encoded by *gpsA*), the terms NAD-dependent and NADP-dependent glycerol-3-phosphate dehydrogenase will be used interchangeably. The NAD-dependent enzyme (EC 1.1.1.8) is encoded, for example, by several genes including *GPD1*, also referred to herein as *DAR1* (coding sequence set forth in SEQ ID NO:1; encoded protein sequence set forth in SEQ ID NO:2), or *GPD2* (coding sequence set forth in SEQ ID NO:3; encoded protein sequence set forth in SEQ ID NO:4), or *GPD3*. The NADP-dependent enzyme (EC 1.1.1.94) is encoded, for example, by *gpsA*.

The terms "glycerol 3-phosphatase", "sn-glycerol 3-phosphatase", "D,L-glycerol phosphatase", and "G3P phosphatase" refer to a polypeptide having an enzymatic activity that is capable of catalyzing the conversion of glycerol 3-phosphate and water to glycerol and inorganic phosphate. G3P phosphatase is encoded, for example, by *GPP1* (coding sequence set forth in SEQ ID NO:5; encoded protein sequence set forth in SEQ ID NO:6), or *GPP2* (coding sequence set forth in SEQ ID NO:7; encoded protein sequence set forth in SEQ ID NO:8).

The term "glycerol dehydratase" or "dehydratase enzyme" refers to a polypeptide having enzyme activity that is capable of catalyzing the conversion of a glycerol molecule to the product, 3-hydroxypropionaldehyde (3-HPA).

For the purposes of the present invention the dehydratase enzymes include a glycerol dehydratase (E.C. 4.2.1.30) and a diol dehydratase (E.C. 4.2.1.28) having preferred substrates of glycerol and 1,2-propanediol, respectively. Genes for dehydratase enzymes have been

identified in *Klebsiella pneumoniae*, *Citrobacter freundii*, *Clostridium pasteurianum*, *Salmonella typhimurium*, *Klebsiella oxytoca*, and *Lactobacillus reuteri*, among others. In each case, the dehydratase is composed of three subunits: the large or "α" subunit, the medium or "β" subunit, and the small or "γ" subunit. The genes are also described in, for example, Daniel *et al.* (*FEMS Microbiol. Rev. 22*, 553 (1999)) and Toraya and Mori (*J. Biol. Chem. 274*, 3372 (1999)). Genes encoding the large or "α" (alpha) subunit of glycerol dehydratase include *dhaB1* (coding sequence set forth in SEQ ID NO:9, encoded protein sequence set forth in SEQ ID NO:10), *gldA* and *dhaB*; genes encoding the medium or "β" (beta) subunit include *dhaB2* (coding sequence set forth in SEQ ID NO:11, encoded protein sequence set forth in SEQ ID NO:12), *gldB*, and *dhaC*; genes encoding the small or "γ" (gamma) subunit include *dhaB3* (coding sequence set forth in SEQ ID NO:13, encoded protein sequence set forth in SEQ ID NO:14), *gldC*, and *dhaE*. Other genes encoding the large or "α" subunit of diol dehydratase include *pduC* and *pddA*; other genes encoding the medium or "β" subunit include *pduD* and *pddB*; and other genes encoding the small or "γ" subunit include *pduE* and *pddC*.

Glycerol and diol dehydratases are subject to mechanism-based suicide inactivation by glycerol and some other substrates (Daniel *et al.*, *FEMS Microbiol. Rev. 22*, 553 (1999)). The term "dehydratase reactivation factor" refers to those proteins responsible for reactivating the dehydratase activity. The terms "dehydratase reactivating activity", "reactivating the dehydratase activity" and "regenerating the dehydratase activity" are used interchangeably and refer to the phenomenon of converting a dehydratase not capable of catalysis of a reaction to one capable of catalysis of a reaction or to the phenomenon of inhibiting the inactivation of a dehydratase or the phenomenon of extending the useful half-life of the dehydratase enzyme *in vivo*. Two proteins have been identified as being involved as the dehydratase

reactivation factor (see, e.g., U.S. Patent No. 6,013,494 and references therein; Daniel *et al.*, *supra*; Toraya and Mori, *J. Biol. Chem. 274*, 3372 (1999); and Tobimatsu *et al.*, *J. Bacteriol. 181*, 4110 (1999)). Genes encoding one of the proteins include, for example, *orfZ*, *dhaB4*, *gdrA*, *pduG* and *ddrA*. Genes encoding the second of the two proteins include, for example, *orfX*, *orf2b*, *gdrB*, *pduH* and *ddrB*.

The terms "1,3-propanediol oxidoreductase", "1,3-propanediol dehydrogenase" and "DhaT" are used interchangeably herein and refer to the polypeptide(s) having an enzymatic activity that is capable of catalyzing the interconversion of 3-HPA and 1,3-propanediol provided the gene(s) encoding such activity is found to be physically or transcriptionally linked to a dehydratase enzyme in its natural (i.e., wild type) setting; for example, the gene is found within a *dha* regulon as is the case with *dhaT* from *Klebsiella pneumoniae*. Genes encoding a 1,3-propanediol oxidoreductase include, but are not limited to, *dhaT* from *Klebsiella pneumoniae*, *Citrobacter freundii*, and *Clostridium pasteurianum*. Each of these genes encode a polypeptide belonging to the family of type III alcohol dehydrogenases, which exhibits a conserved iron-binding motif, and has a preference for the NAD^+/NADH linked interconversion of 3-HPA and 1,3-propanediol (Johnson and Lin, *J. Bacteriol. 169*, 2050 (1987); Daniel *et al.*, *J. Bacteriol. 177*, 2151 (1995); and Leurs *et al.*, *FEMS Microbiol. Lett. 154*, 337 (1997)). Enzymes with similar physical properties have been isolated from *Lactobacillus brevis* and *Lactobacillus buchneri* (Veiga da Dunha and Foster, *Appl. Environ. Microbiol. 58*, 2005 (1992)).

The term "*dha* regulon" refers to a set of associated polynucleotides or open reading frames encoding polypeptides having various biological activities, including but not limited to a dehydratase activity, a reactivation activity, and a 1,3-propanediol oxidoreductase. Typically a *dha* regulon comprises the open reading frames *dhaR*, *orfY*, *dhaT*, *orfX*, *orfW*, *dhaB1*, *dhaB2*, *dhaB3*, and *orfZ* as described in U.S. Patent No. 7,371,558.

The terms "aldehyde dehydrogenase" and "Ald" refer to a polypeptide that catalyzes the conversion of an aldehyde to a carboxylic acid. Aldehyde dehydrogenases may use a redox cofactor such as NAD, NADP, FAD, or PQQ. Typical of aldehyde dehydrogenases is EC 1.2.1.3 (NAD-dependent); EC 1.2.1.4 (NADP-dependent); EC 1.2.99.3 (PQQ-dependent); or EC 1.2.99.7 (FAD-dependent). An example of an NADP-dependent aldehyde dehydrogenase is AldB (SEQ ID NO:16), encoded by the *E. coli* gene *aldB* (coding sequence set forth in SEQ ID NO:15). Examples of NAD-dependent aldehyde dehydrogenases include AldA (SEQ ID NO:18), encoded by the *E. coli* gene *aldA* (coding sequence set forth in SEQ ID NO:17); and AldH (SEQ ID NO:20), encoded by the *E. coli* gene *aldH* (coding sequence set forth in SEQ ID NO:19).

The terms "glucokinase" and "Glk" are used interchangeably herein and refer to a protein that catalyzes the conversion of D-glucose + ATP to glucose 6-phosphate + ADP. Typical of glucokinase is EC 2.7.1.2. Glucokinase is encoded by *glk* in *E. coli*.

The terms "phosphoenolpyruvate carboxylase" and "Ppc" are used interchangeably herein and refer to a protein that catalyzes the conversion of phosphoenolpyruvate + H_2O + CO_2 to phosphate + oxaloacetic acid. Typical of phosphoenolpyruvate carboxylase is EC 4.1.1.31. Phosphoenolpyruvate carboxylase is encoded by *ppc* in *E. coli*.

The terms "glyceraldehyde-3-phosphate dehydrogenase" and "GapA" are used interchangeably herein and refer to a protein having an enzymatic activity capable of catalyzing the conversion of glyceraldehyde 3-phosphate + phosphate + NAD^+ to 3-phospho-D-glyceroyl-phosphate + NADH + H^+. Typical of glyceraldehyde-3-phosphate dehydrogenase is EC 1.2.1.12. Glyceraldehyde-3-phosphate dehydrogenase is encoded by *gapA* in *E. coli*.

The terms "aerobic respiration control protein" and "ArcA" are used interchangeably herein and refer to a global regulatory protein. The aerobic respiration control protein is encoded by *arcA* in *E. coli*.

The terms "methylglyoxal synthase" and "MgsA" are used interchangeably herein and refer to a protein having an enzymatic activity capable of catalyzing the conversion of dihydroxyacetone phosphate to methylglyoxal + phosphate. Typical of methylglyoxal synthase is EC 4.2.3.3. Methylglyoxal synthase is encoded by *mgsA* in *E. coli*.

The terms "phosphogluconate dehydratase" and "Edd" are used interchangeably herein and refer to a protein having an enzymatic activity capable of catalyzing the conversion of 6-phospho-gluconate to 2-keto-3-deoxy-6-phospho-gluconate + H_2O. Typical of phosphogluconate dehydratase is EC 4.2.1.12. Phosphogluconate dehydratase is encoded by *edd* in *E. coli*.

The term "YciK" refers to a putative enzyme encoded by *yciK* which is translationally coupled to *btuR*, the gene encoding Cob(I)alamin adenosyltransferase in *E. coli*.

The term "cob(I)alamin adenosyltransferase" refers to an enzyme capable of transferring a deoxyadenosyl moiety from ATP to the reduced corrinoid. Typical of cob(I)alamin adenosyltransferase is EC 2.5.1.17. Cob(I)alamin adenosyltransferase is encoded by the gene "*btuR*" in *E. coli*, "*cobA*" in *Salmonella typhimurium*, and "*cobO*" in *Pseudomonas denitrificans*.

The terms "galactose-proton symporter" and "GalP" are used interchangeably herein and refer to a protein having an enzymatic activity capable of transporting a sugar and a proton from the periplasm to the cytoplasm. D-glucose is a preferred substrate for GalP. Galactose-proton symporter is encoded by *galP* in *Escherichia coli* (coding sequence set forth in SEQ ID NO:21, encoded protein sequence set forth in SEQ ID NO:22).

The term "non-specific catalytic activity" refers to the polypeptide(s) having an enzymatic activity capable of catalyzing the interconversion of 3-HPA and 1,3-propanediol and specifically excludes 1,3-propanediol oxidoreductase(s). Typically these enzymes are alcohol dehydrogenases. Such enzymes may utilize cofactors other than

NAD^+/NADH, including but not limited to flavins such as FAD or FMN. A gene for a non-specific alcohol dehydrogenase (*yqhD*) is found, for example, to be endogenously encoded and functionally expressed within *E. coli* K-12 strains.

The terms "1.6 long GI promoter", "1.20 short/long GI Promoter", "1.5 long GI promoter", "P1.6", "P1.5" and "P1.20" refer to polynucleotides or fragments containing a promoter from the *Streptomyces lividans* glucose isomerase gene as described in U.S. Patent No. 7,132,527. These promoter fragments include a mutation which decreases their activities as compared to the wild type *Streptomyces lividans* glucose isomerase gene promoter.

The terms "function" and "enzyme function" are used interchangeably herein and refer to the catalytic activity of an enzyme in altering the rate at which a specific chemical reaction occurs without itself being consumed by the reaction. It is understood that such an activity may apply to a reaction in equilibrium where the production of either product or substrate may be accomplished under suitable conditions.

The terms "polypeptide" and "protein" are used interchangeably herein.

The terms "carbon substrate" and "carbon source" are used interchangeably herein and refer to a carbon source capable of being metabolized by the recombinant bacteria disclosed herein and, particularly, carbon sources comprising sucrose. The carbon source may further comprise other monosaccharides, disaccharides, oligosaccharides; or polysaccharides.

The terms "host cell" and "host bacterium" are used interchangeably herein and refer to a bacterium capable of receiving foreign or heterologous genes and capable of expressing those genes to produce an active gene product.

The term "production microorganism" as used herein refers to a microorganism, including, but not limited to, those that are recombinant,

used to make a specific product such as 1,3-propanediol, glycerol, 3-hydroxypropionic acid, polyunsaturated fatty acids, and the like.

As used herein, "nucleic acid" means a polynucleotide and includes a single or double-stranded polymer of deoxyribonucleotide or ribonucleotide bases. Nucleic acids may also include fragments and modified nucleotides. Thus, the terms "polynucleotide", "nucleic acid sequence", "nucleotide sequence" or "nucleic acid fragment" are used interchangeably herein and refer to a polymer of RNA or DNA that is single- or double-stranded, optionally containing synthetic, non-natural or altered nucleotide bases. Nucleotides (usually found in their 5'-monophosphate form) are referred to by their single letter designation as follows: "A" for adenylate or deoxyadenylate (for RNA or DNA, respectively), "C" for cytidylate or deoxycytidylate, "G" for guanylate or deoxyguanylate, "U" for uridylate, "T" for deoxythymidylate, "R" for purines (A or G), "Y" for pyrimidines (C or T), "K" for G or T, "H" for A or C or T, "I" for inosine, and "N" for any nucleotide.

A polynucleotide may be a polymer of RNA or DNA that is single- or double-stranded, that optionally contains synthetic, non-natural or altered nucleotide bases. A polynucleotide in the form of a polymer of DNA may be comprised of one or more segments of cDNA, genomic DNA, synthetic DNA, or mixtures thereof.

"Gene" refers to a nucleic acid fragment that expresses a specific protein, and which may refer to the coding region alone or may include regulatory sequences preceding (5' non-coding sequences) and following (3' non-coding sequences) the coding sequence. "Native gene" refers to a gene as found in nature with its own regulatory sequences. "Chimeric gene" refers to any gene that is not a native gene, comprising regulatory and coding sequences that are not found together in nature. Accordingly, a chimeric gene may comprise regulatory sequences and coding sequences that are derived from different sources, or regulatory sequences and coding sequences derived from the same source, but arranged in a manner different than

that found in nature. "Endogenous gene" refers to a native gene in its natural location in the genome of an organism. A "foreign" gene refers to a gene that is introduced into the host organism by gene transfer. Foreign genes can comprise genes inserted into a non-native organism, genes introduced into a new location within the native host, or chimeric genes.

The term "native nucleotide sequence" refers to a nucleotide sequence that is normally found in the host microorganism.

The term "non-native nucleotide sequence" refers to a nucleotide sequence that is not normally found in the host microorganism.

The term "native polypeptide" refers to a polypeptide that is normally found in the host microorganism.

The term "non-native polypeptide" refers to a polypeptide that is not normally found in the host microorganism.

The terms "encoding" and "coding" are used interchangeably herein and refer to the process by which a gene, through the mechanisms of transcription and translation, produces an amino acid sequence.

The term "coding sequence" refers to a nucleotide sequence that codes for a specific amino acid sequence.

"Suitable regulatory sequences" refer to nucleotide sequences located upstream (5' non-coding sequences), within, or downstream (3' non-coding sequences) of a coding sequence, and which influence the transcription, RNA processing or stability, or translation of the associated coding sequence. Regulatory sequences may include promoters, enhancers, silencers, 5' untranslated leader sequence (e.g., between the transcription start site and the translation initiation codon), introns, polyadenylation recognition sequences, RNA processing sites, effector binding sites and stem-loop structures.

The term "expression cassette" refers to a fragment of DNA comprising the coding sequence of a selected gene and regulatory sequences preceding (5' non-coding sequences) and following (3' non-coding sequences) the coding sequence that are required for expression

of the selected gene product. Thus, an expression cassette is typically composed of: 1) a promoter sequence; 2) a coding sequence (i.e., ORF) and, 3) a 3' untranslated region (e.g., a terminator) that, in eukaryotes, usually contains a polyadenylation site. The expression cassette(s) is usually included within a vector, to facilitate cloning and transformation. Different organisms, including bacteria, yeast, and fungi, can be transformed with different expression cassettes as long as the correct regulatory sequences are used for each host.

"Transformation" refers to the transfer of a nucleic acid molecule into a host organism, resulting in genetically stable inheritance. The nucleic acid molecule may be a plasmid that replicates autonomously, for example, or it may integrate into the genome of the host organism. Host organisms transformed with the nucleic acid fragments are referred to as "recombinant" or "transformed" organisms or "transformants". "Stable transformation" refers to the transfer of a nucleic acid fragment into a genome of a host organism, including both nuclear and organellar genomes, resulting in genetically stable inheritance. In contrast, "transient transformation" refers to the transfer of a nucleic acid fragment into the nucleus, or DNA-containing organelle, of a host organism resulting in gene expression without integration or stable inheritance.

"Codon degeneracy" refers to the nature in the genetic code permitting variation of the nucleotide sequence without affecting the amino acid sequence of an encoded polypeptide. The skilled artisan is well aware of the "codon-bias" exhibited by a specific host cell in usage of nucleotide codons to specify a given amino acid. Therefore, when synthesizing a gene for improved expression in a host cell, it is desirable to design the gene such that its frequency of codon usage approaches the frequency of preferred codon usage of the host cell.

The terms "subfragment that is functionally equivalent" and "functionally equivalent subfragment" are used interchangeably herein. These terms refer to a portion or subsequence of an isolated nucleic acid fragment in which the ability to alter gene expression or produce a certain

phenotype is retained whether or not the fragment or subfragment encodes an active enzyme. Chimeric genes can be designed for use in suppression by linking a nucleic acid fragment or subfragment thereof, whether or not it encodes an active enzyme, in the sense or antisense orientation relative to a promoter sequence.

The term "conserved domain" or "motif" means a set of amino acids conserved at specific positions along an aligned sequence of evolutionarily related proteins. While amino acids at other positions can vary between homologous proteins, amino acids that are highly conserved at specific positions indicate amino acids that are essential in the structure, the stability, or the activity of a protein.

The terms "substantially similar" and "corresponds substantially" are used interchangeably herein. They refer to nucleic acid fragments wherein changes in one or more nucleotide bases do not affect the ability of the nucleic acid fragment to mediate gene expression or produce a certain phenotype. These terms also refer to modifications of the nucleic acid fragments of the instant invention such as deletion or insertion of one or more nucleotides that do not substantially alter the functional properties of the resulting nucleic acid fragment relative to the initial, unmodified fragment. It is therefore understood, as those skilled in the art will appreciate, that the invention encompasses more than the specific exemplary sequences. Moreover, the skilled artisan recognizes that substantially similar nucleic acid sequences encompassed by this invention are also defined by their ability to hybridize (under moderately stringent conditions, e.g., 0.5X SSC (standard sodium citrate), 0.1% SDS (sodium dodecyl sulfate), 60 °C) with the sequences exemplified herein, or to any portion of the nucleotide sequences disclosed herein and which are functionally equivalent to any of the nucleic acid sequences disclosed herein. Stringency conditions can be adjusted to screen for moderately similar fragments, such as homologous sequences from distantly related organisms, to highly similar fragments, such as genes that duplicate functional enzymes from closely related organisms. Post-hybridization

washes determine stringency conditions.

The term "selectively hybridizes" includes reference to hybridization, under stringent hybridization conditions, of a nucleic acid sequence to a specified nucleic acid target sequence to a detectably greater degree (e.g., at least 2-fold over background) than its hybridization to non-target nucleic acid sequences and to the substantial exclusion of non-target nucleic acids. Selectively hybridizing sequences are two nucleotide sequences wherein the complement of one of the nucleotide sequences typically has about at least 80% sequence identity, or 90% sequence identity, up to and including 100% sequence identity (i.e., fully complementary) to the other nucleotide sequence.

The term "stringent conditions" or "stringent hybridization conditions" includes reference to conditions under which a probe will selectively hybridize to its target sequence. Probes are typically single stranded nucleic acid sequences which are complementary to the nucleic acid sequences to be detected. Probes are "hybridizable" to the nucleic acid sequence to be detected. Generally, a probe is less than about 1000 nucleotides in length, optionally less than 500 nucleotides in length.

Hybridization methods are well defined. Typically the probe and sample are mixed under conditions which will permit nucleic acid hybridization. This involves contacting the probe and sample in the presence of an inorganic or organic salt under the proper concentration and temperature conditions. Optionally a chaotropic agent may be added. Nucleic acid hybridization is adaptable to a variety of assay formats. One of the most suitable is the sandwich assay format. A primary component of a sandwich-type assay is a solid support. The solid support has adsorbed to it or covalently coupled to it an immobilized nucleic acid probe that is unlabeled and complementary to one portion of the sequence.

Stringent conditions are sequence-dependent and will be different in different circumstances. By controlling the stringency of the hybridization and/or washing conditions, target sequences can be identified which are 100% complementary to the probe (homologous

probing). Alternatively, stringency conditions can be adjusted to allow some mismatching in sequences so that lower degrees of similarity are detected (heterologous probing).

Typically, stringent conditions will be those in which the salt concentration is less than about 1.5 M Na ion, typically about 0.01 to 1.0 M Na ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30 °C for short probes (e.g., 10 to 50 nucleotides) and at least about 60 °C for long probes (e.g., greater than 50 nucleotides). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide. Exemplary low stringency conditions include hybridization with a buffer solution of 30 to 35% formamide, 1 M NaCl, 1% SDS (sodium dodecyl sulfate) at 37 °C, and a wash in 1X to 2X SSC (20X SSC = 3.0 M NaCl/0.3 M trisodium citrate) at 50 to 55 °C. Exemplary moderate stringency conditions include hybridization in 40 to 45% formamide, 1 M NaCl, 1% SDS at 37 °C, and a wash in 0.5X to 1X SSC at 55 to 60 °C. Exemplary high stringency conditions include hybridization in 50% formamide, 1 M NaCl, 1% SDS at 37 °C, and a wash in 0.1X SSC at 60 to 65 °C.

Specificity is typically the function of post-hybridization washes, the critical factors being the ionic strength and temperature of the final wash solution. For DNA-DNA hybrids, the thermal melting point (T_m) can be approximated from the equation of Meinkoth et al., *Anal. Biochem.* 138:267-284 (1984): $T_m = 81.5\ °C + 16.6\ (\log M) + 0.41\ (\%GC) - 0.61\ (\%\ form) - 500/L$; where M is the molarity of monovalent cations, %GC is the percentage of guanosine and cytosine nucleotides in the DNA, % form is the percentage of formamide in the hybridization solution, and L is the length of the hybrid in base pairs. The T_m is the temperature (under defined ionic strength and pH) at which 50% of a complementary target sequence hybridizes to a perfectly matched probe. T_m is reduced by about 1 °C for each 1% of mismatching; thus, T_m, hybridization and/or wash conditions can be adjusted to hybridize to sequences of the desired identity. For example, if sequences with $\geq$90% identity are sought, the T_m

can be decreased 10 °C. Generally, stringent conditions are selected to be about 5 °C lower than T_m for the specific sequence and its complement at a defined ionic strength and pH. However, severely stringent conditions can utilize a hybridization and/or wash at 1, 2, 3, or 4 °C lower than the T_m; moderately stringent conditions can utilize a hybridization and/or wash at 6, 7, 8, 9, or 10 °C lower than the T_m; low stringency conditions can utilize a hybridization and/or wash at 11, 12, 13, 14, 15, or 20 °C lower than the T_m. Using the equation, hybridization and wash compositions, and desired T_m, those of ordinary skill will understand that variations in the stringency of hybridization and/or wash solutions are inherently described. If the desired degree of mismatching results in a T_m of less than 45 °C (aqueous solution) or 32 °C (formamide solution) it is preferred to increase the SSC concentration so that a higher temperature can be used. An extensive guide to the hybridization of nucleic acids is found in Tijssen, *Laboratory Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Acid Probes*, Part I, Chapter 2 "Overview of principles of hybridization and the strategy of nucleic acid probe assays", Elsevier, New York (1993); and *Current Protocols in Molecular Biology*, Chapter 2, Ausubel et al., Eds., Greene Publishing and Wiley-Interscience, New York (1995). Hybridization and/or wash conditions can be applied for at least 10, 30, 60, 90, 120, or 240 minutes.

"Sequence identity" or "identity" in the context of nucleic acid or polypeptide sequences refers to the nucleic acid bases or amino acid residues in two sequences that are the same when aligned for maximum correspondence over a specified comparison window.

Thus, "percentage of sequence identity" refers to the value determined by comparing two optimally aligned sequences over a comparison window, wherein the portion of the polynucleotide or polypeptide sequence in the comparison window may comprise additions or deletions (i.e., gaps) as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of

positions at which the identical nucleic acid base or amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison and multiplying the results by 100 to yield the percentage of sequence identity. Useful examples of percent sequence identities include, but are not limited to, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95%, or any integer percentage from 50% to 100%. These identities can be determined using any of the programs described herein.

Sequence alignments and percent identity or similarity calculations may be determined using a variety of comparison methods designed to detect homologous sequences including, but not limited to, the MegAlign™ program of the LASERGENE bioinformatics computing suite (DNASTAR Inc., Madison, WI). Within the context of this application it will be understood that where sequence analysis software is used for analysis, that the results of the analysis will be based on the "default values" of the program referenced, unless otherwise specified. As used herein "default values" will mean any set of values or parameters that originally load with the software when first initialized.

The "Clustal V method of alignment" corresponds to the alignment method labeled Clustal V (described by Higgins and Sharp, *CABIOS.* 5:151-153 (1989); Higgins, D.G. et al., *Comput. Appl. Biosci. 8*:189-191 (1992)) and found in the MegAlign™ program of the LASERGENE bioinformatics computing suite (DNASTAR Inc., Madison, WI). For multiple alignments, the default values correspond to GAP PENALTY=10 and GAP LENGTH PENALTY=10. Default parameters for pairwise alignments and calculation of percent identity of protein sequences using the Clustal V method are KTUPLE=1, GAP PENALTY=3, WINDOW=5 and DIAGONALS SAVED=5. For nucleic acids these parameters are KTUPLE=2, GAP PENALTY=5, WINDOW=4 and DIAGONALS SAVED=4. After alignment of the sequences using the Clustal V program, it is possible to obtain a "percent identity" by viewing the "sequence distances"

table in the same program.

The "Clustal W method of alignment" corresponds to the alignment method labeled Clustal W (described by Higgins and Sharp, *supra*; Higgins, D.G. et al., *supra*) and found in the MegAlign™ v6.1 program of the LASERGENE bioinformatics computing suite (DNASTAR Inc., Madison, WI). Default parameters for multiple alignment correspond to GAP PENALTY=10, GAP LENGTH PENALTY=0.2, Delay Divergen Seqs(%)=30, DNA Transition Weight=0.5, Protein Weight Matrix=Gonnet Series, DNA Weight Matrix=IUB. After alignment of the sequences using the Clustal W program, it is possible to obtain a "percent identity" by viewing the "sequence distances" table in the same program.

"BLASTN method of alignment" is an algorithm provided by the National Center for Biotechnology Information (NCBI) to compare nucleotide sequences using default parameters. The "BLASTP method of alignment" is an algorithm provided by the NCBI to compare protein sequences using default parameters.

It is well understood by one skilled in the art that many levels of sequence identity are useful in identifying polypeptides, from other species, wherein such polypeptides have the same or similar function or activity. Useful examples of percent identities include, but are not limited to, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95%, or any integer percentage from 50% to 100%. Indeed, any integer amino acid identity from 50% to 100% may be useful in describing the present invention, such as 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99%. Also, of interest is any full-length or partial complement of this isolated nucleotide fragment.

Thus, the invention encompasses more than the specific exemplary nucleotide sequences disclosed herein. For example, alterations in the gene sequence which reflect the degeneracy of the genetic code are

contemplated. Also, it is well known in the art that alterations in a gene which result in the production of a chemically equivalent amino acid at a given site, but do not affect the functional properties of the encoded protein are common. Substitutions are defined for the discussion herein as exchanges within one of the following five groups:

1. Small aliphatic, nonpolar or slightly polar residues: Ala, Ser, Thr (Pro, Gly);
2. Polar, negatively charged residues and their amides: Asp, Asn, Glu, Gln;
3. Polar, positively charged residues: His, Arg, Lys;
4. Large aliphatic, nonpolar residues: Met, Leu, Ile, Val (Cys); and
5. Large aromatic residues: Phe, Tyr, Trp.

Thus, a codon for the amino acid alanine, a hydrophobic amino acid, may be substituted by a codon encoding another less hydrophobic residue (such as glycine) or a more hydrophobic residue (such as valine, leucine, or isoleucine). Similarly, changes which result in substitution of one negatively charged residue for another (such as aspartic acid for glutamic acid) or one positively charged residue for another (such as lysine for arginine) can also be expected to produce a functionally equivalent product. In many cases, nucleotide changes which result in alteration of the N-terminal and C-terminal portions of the protein molecule would also not be expected to alter the activity of the protein.

Each of the proposed modifications is well within the routine skill in the art, as is determination of retention of biological activity of the encoded products. Moreover, the skilled artisan recognizes that substantially similar sequences encompassed by this invention are also defined by their ability to hybridize under stringent conditions, as defined above.

Preferred substantially similar nucleic acid fragments of the instant invention are those nucleic acid fragments whose nucleotide sequences are at least 70% identical to the nucleotide sequence of the nucleic acid fragments reported herein. More preferred nucleic acid fragments are at

least 90% identical to the nucleotide sequence of the nucleic acid fragments reported herein. Most preferred are nucleic acid fragments that are at least 95% identical to the nucleotide sequence of the nucleic acid fragments reported herein.

A "substantial portion" of an amino acid or nucleotide sequence is that portion comprising enough of the amino acid sequence of a polypeptide or the nucleotide sequence of a gene to putatively identify that polypeptide or gene, either by manual evaluation of the sequence by one skilled in the art, or by computer-automated sequence comparison and identification using algorithms such as BLAST (Basic Local Alignment Search Tool; Altschul, S. F., et al., *J. Mol. Biol.*, 215:403-410 (1993)). In general, a sequence of ten or more contiguous amino acids or thirty or more nucleotides is necessary in order to putatively identify a polypeptide or nucleic acid sequence as homologous to a known protein or gene. Moreover, with respect to nucleotide sequences, gene-specific oligonucleotide probes comprising 20-30 contiguous nucleotides may be used in sequence-dependent methods of gene identification (e.g., Southern hybridization) and isolation (e.g., *in situ* hybridization of bacterial colonies or bacteriophage plaques). In addition, short oligonucleotides of 12-15 bases may be used as amplification primers in PCR in order to obtain a particular nucleic acid fragment comprising the primers. Accordingly, a "substantial portion" of a nucleotide sequence comprises enough of the sequence to specifically identify and/or isolate a nucleic acid fragment comprising the sequence. The instant specification teaches the complete amino acid and nucleotide sequence encoding particular proteins. The skilled artisan, having the benefit of the sequences as reported herein, may now use all or a substantial portion of the disclosed sequences for purposes known to those skilled in this art.

The term "complementary" describes the relationship between two sequences of nucleotide bases that are capable of Watson-Crick base-pairing when aligned in an anti-parallel orientation. For example, with respect to DNA, adenosine is capable of base-pairing with thymine and

cytosine is capable of base-pairing with guanine. Accordingly, the instant invention may make use of isolated nucleic acid molecules that are complementary to the complete sequences as reported in the accompanying Sequence Listing and the specification as well as those substantially similar nucleic acid sequences.

The term "isolated" refers to a polypeptide or nucleotide sequence that is removed from at least one component with which it is naturally associated.

"Promoter" refers to a DNA sequence capable of controlling the expression of a coding sequence or functional RNA. The promoter sequence consists of proximal and more distal upstream elements, the latter elements often referred to as enhancers. Accordingly, an "enhancer" is a DNA sequence that can stimulate promoter activity, and may be an innate element of the promoter or a heterologous element inserted to enhance the level or tissue-specificity of a promoter. Promoters may be derived in their entirety from a native gene, or be composed of different elements derived from different promoters found in nature, or even comprise synthetic DNA segments. It is understood by those skilled in the art that different promoters may direct the expression of a gene in different tissues or cell types, or at different stages of development, or in response to different environmental conditions. It is further recognized that since in most cases the exact boundaries of regulatory sequences have not been completely defined, DNA fragments of some variation may have identical promoter activity. Promoters that cause a gene to be expressed in most cell types at most times are commonly referred to as "constitutive promoters".

"3' non-coding sequences", "transcription terminator" and "termination sequences" are used interchangeably herein and refer to DNA sequences located downstream of a coding sequence, including polyadenylation recognition sequences and other sequences encoding regulatory signals capable of affecting mRNA processing or gene expression. The polyadenylation signal is usually characterized by

affecting the addition of polyadenylic acid tracts to the 3' end of the mRNA precursor.

The term "operably linked" refers to the association of nucleic acid sequences on a single nucleic acid fragment so that the function of one is affected by the other. For example, a promoter is operably linked with a coding sequence when it is capable of affecting the expression of that coding sequence (i.e., the coding sequence is under the transcriptional control of the promoter). Coding sequences can be operably linked to regulatory sequences in a sense or antisense orientation. In another example, the complementary RNA regions of the invention can be operably linked, either directly or indirectly, 5' to the target mRNA, or 3' to the target mRNA, or within the target mRNA, or a first complementary region is 5' and its complement is 3' to the target mRNA.

Standard recombinant DNA and molecular cloning techniques used herein are well known in the art and are described more fully in Sambrook, J., Fritsch, E.F. and Maniatis, T. Molecular Cloning: A Laboratory Manual; Cold Spring Harbor Laboratory: Cold Spring Harbor, NY (1989). Transformation methods are well known to those skilled in the art and are described *infra*.

"PCR" or "polymerase chain reaction" is a technique for the synthesis of large quantities of specific DNA segments and consists of a series of repetitive cycles (Perkin Elmer Cetus Instruments, Norwalk, CT). Typically, the double-stranded DNA is heat denatured, the two primers complementary to the 3' boundaries of the target segment are annealed at low temperature and then extended at an intermediate temperature. One set of these three consecutive steps is referred to as a "cycle".

A "plasmid" or "vector" is an extra chromosomal element often carrying genes that are not part of the central metabolism of the cell, and usually in the form of circular double-stranded DNA fragments. Such elements may be autonomously replicating sequences, genome integrating sequences, phage or nucleotide sequences, linear or circular, of a single- or double-stranded DNA or RNA, derived from any source, in

which a number of nucleotide sequences have been joined or recombined into a unique construction which is capable of introducing an expression cassette(s) into a cell.

The term "genetically altered" refers to the process of changing hereditary material by genetic engineering, transformation and/or mutation.

The term "recombinant" refers to an artificial combination of two otherwise separated segments of sequence, e.g., by chemical synthesis or by the manipulation of isolated segments of nucleic acids by genetic engineering techniques. "Recombinant" also includes reference to a cell or vector, that has been modified by the introduction of a heterologous nucleic acid or a cell derived from a cell so modified, but does not encompass the alteration of the cell or vector by naturally occurring events (e.g., spontaneous mutation, natural transformation, natural transduction, natural transposition) such as those occurring without deliberate human intervention.

The terms "recombinant construct", "expression construct", "chimeric construct", "construct", and "recombinant DNA construct", are used interchangeably herein. A recombinant construct comprises an artificial combination of nucleic acid fragments, e.g., regulatory and coding sequences that are not found together in nature. For example, a recombinant construct may comprise regulatory sequences and coding sequences that are derived from different sources, or regulatory sequences and coding sequences derived from the same source, but arranged in a manner different than that found in nature. Such a construct may be used by itself or may be used in conjunction with a vector. If a vector is used, then the choice of vector is dependent upon the method that will be used to transform host cells as is well known to those skilled in the art. For example, a plasmid vector can be used. The skilled artisan is well aware of the genetic elements that must be present on the vector in order to successfully transform, select and propagate host cells comprising any of the isolated nucleic acid fragments of the invention.

The skilled artisan will also recognize that different independent transformation events may result in different levels and patterns of expression (Jones et al., *EMBO J.* 4:2411-2418 (1985); De Almeida et al., *Mol. Gen. Genetics* 218:78-86 (1989)), and thus that multiple events may need be screened in order to obtain lines displaying the desired expression level and pattern. Such screening may be accomplished by Southern analysis of DNA, Northern analysis of mRNA expression, immunoblotting analysis of protein expression, or phenotypic analysis, among others.

The term "expression", as used herein, refers to the production of a functional end-product (e.g., an mRNA or a protein [either precursor or mature]).

The term "introduced" means providing a nucleic acid (e.g., expression construct) or protein into a cell. Introduced includes reference to the incorporation of a nucleic acid into a eukaryotic or prokaryotic cell where the nucleic acid may be incorporated into the genome of the cell, and includes reference to the transient provision of a nucleic acid or protein to the cell. Introduced includes reference to stable or transient transformation methods, as well as sexually crossing. Thus, "introduced" in the context of inserting a nucleic acid fragment (e.g., a recombinant construct/expression construct) into a cell, means "transfection" or "transformation" or "transduction" and includes reference to the incorporation of a nucleic acid fragment into a eukaryotic or prokaryotic cell where the nucleic acid fragment may be incorporated into the genome of the cell (e.g., chromosome, plasmid, plastid or mitochondrial DNA), converted into an autonomous replicon, or transiently expressed (e.g., transfected mRNA).

The term "homologous" refers to proteins or polypeptides of common evolutionary origin with similar catalytic function. The invention may include bacteria producing homologous proteins via recombinant technology.

Disclosed herein are variant sucrose transporter polypeptides that enable bacterial growth over a wide range of gene expression levels and sucrose concentrations. Sucrose transporter polypeptides are polypeptides that are capable of mediating the transport of sucrose into microbial cells. Sucrose transporters known in the art, such as CscB from *E. coli*, function as H^+/sucrose symporters, which transport one proton for every sucrose molecule transported, thereby coupling the energy of the proton motive force to sucrose transport. Such active transport allows accumulation of sucrose against a concentration gradient. Mutations which change certain amino acids in CscB that result in polypeptides unable to catalyze active uptake of sucrose, but are able to catalyze equilibrium exchange across a membrane have been described by Vadyvaloo et al. (*J. Mol. Biol.* 358:1051-1059, 2006). The sucrose transporter polypeptides disclosed herein are novel variants that have lost the ability to actively transport sucrose into microbial cells against a concentration gradient, but have the ability to transport sucrose by facilitated diffusion. These variant sucrose transporter polypeptides also enable faster sucrose utilization in bacteria than the native CscB transporter polypeptide. Sucrose transport by facilitated diffusion mitigates the toxicity associated with excess sucrose uptake because sucrose will not accumulate within the cells to concentrations that are higher than extracellular levels. Therefore, microbial cells having sucrose transport by facilitated diffusion are able to grow over a wider range of sucrose concentrations than cells having active sucrose transport.

In some embodiments, the sucrose transporter polypeptides disclosed herein are variants of the wild-type sucrose transporter polypeptide CscB from *E. coli* ATCC®13281 (set forth in SEQ ID NO:26, nucleotide coding sequence set forth in SEQ ID NO:25). These sucrose transporter polypeptides have an amino acid change from arginine to alanine at amino acid position 300, i.e., R300A mutation, (SEQ ID NO:100, nucleotide coding sequence set forth in SEQ ID NO:99) or an amino acid change from arginine to leucine at amino acid position 300, i.e., R300L

mutation, (SEQ ID NO:102, nucleotide coding sequence set forth in SEQ ID NO:101) and at least one other mutation which results in faster sucrose utilization, as described by Chen et al. (U.S. Patent Application No. 13/210488, filed August 16, 2011), i.e., either an amino acid change or a truncation of the amino acid sequence. Accordingly, in these embodiments, the variant sucrose transporter polypeptides have: an amino acid sequence that has at least 95% identity to an amino acid sequence as set forth in SEQ ID NO:26 based on a Clustal W method of alignment and have an amino acid change from arginine to alanine or arginine to leucine at position 300, and comprise:

(a) at least one amino acid change selected from the group consisting of:

(i) glutamine to histidine at position 353

(ii) leucine to proline at position 61;

(iii) phenylalanine to leucine at position 159;

(iv) glycine to cysteine at position 162;

(v) proline to histidine at position 169;

(vi) leucine to tryptophan at position 61;

(vii) leucine to histidine at position 61;

(viii) leucine to phenylalanine at position 61; and

(ix) leucine to tyrosine at position 61; or

(b) a length of 402 to 407 amino acids from the N-terminus; or

(c) a length of 402 to 407 amino acids from the N-terminus, and at least one of the amino acid changes of (a).

In some embodiments, the sucrose transporter polypeptides are variants of sucrose transporter polypeptides from various sources (see Table 1), having an amino acid change to alanine or leucine at a position equivalent to amino acid position 300 when compared with a reference amino acid sequence of CscB (SEQ ID NO:26). The corresponding amino acid positions in the various sucrose transporter polypeptides, relative to the reference amino acid sequence, can be readily determined by one skilled in the art using sequence alignment algorithms, such as Clustal W,

Clustal V, and BLASTP, which are described above. Accordingly, in these embodiments, the variant sucrose transporter polypeptides have an amino acid sequence that has at least 95% identity based on a Clustal W method of alignment to an amino acid sequence selected from the group consisting of SEQ ID NOs:68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, and 98, and an amino acid at an equivalent position when compared with a reference amino acid sequence of CscB (SEQ ID NO:26) selected from the group consisting of:

(a) alanine at a position equivalent to position 300; and

(b) leucine at a position equivalent to position 300;

In some embodiments, the sucrose transporter polypeptides are variants of sucrose transporter polypeptides from various sources (see Table 1) having an amino acid change to alanine or leucine at a position equivalent to amino acid position 300 when compared with a reference amino acid sequence of CscB (SEQ ID NO:26), as described above, and further comprise:

(a) at least one of the following amino acids at an equivalent position when compared with the reference amino acid sequence of SEQ ID NO:26:

(i) histidine at a position equivalent to position 353;

(ii) proline at a position equivalent to position 61;

(iii) leucine at a position equivalent to position 159;

(iv) cysteine at a position equivalent to position 162;

(v) histidine at a position equivalent to position 169;

(vi) tryptophan at a position equivalent to position 61;

(vii) histidine at a position equivalent to position 61;

(viii) phenylalanine at a position equivalent to position 61;

(ix) tyrosine at a position equivalent to position 61; and/or

(b) truncation at a position equivalent to position 407, 406, 405, 404, 403, or 402 when compared with the reference amino acid sequence of SEQ ID NO:26.

In some embodiments, the variant sucrose transporter polypeptides

have an amino acid sequence selected from the group consisting of: SEQ ID NOs:100, 102, 104, 106, 108, 110, and 112.

Also disclosed herein are bacteria comprising in their genome or on at least one recombinant construct a nucleotide sequence encoding a variant sucrose transporter polypeptide and a nucleotide sequence encoding a polypeptide having sucrose hydrolase activity. The nucleotide sequences are each operably linked to the same or a different promoter. These bacteria are able to grow over a wider range of gene expression levels and sucrose concentrations than bacteria having native sucrose transporter polypeptides which actively transport sucrose. Accordingly, in these embodiments, the recombinant bacteria comprise in their genome or on at least one recombinant construct:

(a) a nucleotide sequence encoding a variant sucrose transporter polypeptide having an amino acid sequence that has at least 95% identity based on a Clustal W method of alignment to an amino acid sequence selected from the group consisting of SEQ ID NOs:26, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, and 98, and an amino acid at an equivalent position when compared with a reference amino acid sequence of SEQ ID NO:26 selected from the group consisting of:

(i) alanine at a position equivalent to position 300; and

(ii) leucine at a position equivalent to position 300; and

(b) a nucleotide sequence encoding a polypeptide having sucrose hydrolase activity;

wherein (a) and (b) are each operably linked to the same or a different promoter, further wherein the recombinant bacteria are capable of metabolizing sucrose.

In some embodiments, the recombinant bacteria comprise a variant sucrose transporter polypeptide which further comprises:

(a) at least one of the following amino acids at an equivalent position when compared with the reference amino acid sequence of SEQ ID NO:26:

(i) histidine at a position equivalent to position 353;

(ii) proline at a position equivalent to position 61;

(iii) leucine at a position equivalent to position 159;

(iv) cysteine at a position equivalent to position 162;

(v) histidine at a position equivalent to position 169;

(vi) tryptophan at a position equivalent to position 61;

(vii) histidine at a position equivalent to position 61;

(viii) phenylalanine at a position equivalent to position 61;

(ix) tyrosine at a position equivalent to position 61; and/or

(b) truncation at a position equivalent to position 407, 406, 405, 404, 403, or 402 when compared with the reference amino acid sequence of SEQ ID NO:26.

Recombinant bacteria comprising a nucleotide sequence encoding a variant sucrose transporter polypeptide, as described above, and a nucleotide sequence encoding a polypeptide having sucrose hydrolase activity may be constructed by introducing the nucleotide sequences into a suitable host bacterium, either into the genome or on at least one recombinant construct, using methods known in the art, as described below. In some embodiments, the recombinant bacteria are capable of metabolizing sucrose to produce glycerol and/or glycerol-derived products.

Suitable host bacteria for use in the construction of the recombinant bacteria disclosed herein include, but are not limited to, organisms of the genera: *Escherichia*, *Streptococcus*, *Agrobacterium*, *Bacillus*, *Corynebacterium*, *Lactobacillus*, *Clostridium*, *Gluconobacter*, *Citrobacter*, *Enterobacter*, *Klebsiella*, *Aerobacter*, *Methylobacter*, *Salmonella*, *Streptomyces*, and *Pseudomonas*.

In some embodiments, the host bacterium is selected from the genera: *Escherichia*, *Klebsiella*, *Citrobacter*, and *Aerobacter*.

In some embodiments, the host bacterium is *Escherichia coli*.

In some embodiments, the host bacterium is PTS minus. In these embodiments, the host bacterium is PTS minus in its native state, or may be rendered PTS minus through inactivation of a PTS gene as described

below.

In production microorganisms, it is sometimes desirable to unlink the transport of sugars and the use of phosphoenolpyruvate (PEP) for phosphorylation of the sugars being transported.

The term "down-regulated" refers to reduction in, or abolishment of, the activity of active protein(s), as compared to the activity of the wild-type protein(s). The PTS may be inactivated (resulting in a "PTS minus" organism) by down-regulating expression of one or more of the endogenous genes encoding the proteins required in this type of transport. Down-regulation typically occurs when one or more of these genes has a "disruption", referring to an insertion, deletion, or targeted mutation within a portion of that gene, that results in either a complete gene knockout such that the gene is deleted from the genome and no protein is translated or a protein has been translated such that it has an insertion, deletion, amino acid substitution or other targeted mutation. The location of the disruption in the protein may be, for example, within the N-terminal portion of the protein or within the C-terminal portion of the protein. The disrupted protein will have impaired activity with respect to the protein that was not disrupted, and can be non-functional. Down-regulation that results in low or lack of expression of the protein, could also result via manipulating the regulatory sequences, transcription and translation factors and/or signal transduction pathways or by use of sense, antisense or RNAi technology, or similar mechanisms known to skilled artisans.

The recombinant bacteria disclosed herein comprise in their genome or on at least one recombinant construct, a nucleotide sequence encoding a polypeptide having sucrose hydrolase activity. Polypeptides having sucrose hydrolase activity have the ability to catalyze the hydrolysis of sucrose to produce fructose and glucose. Polypeptides having sucrose hydrolase activity are known, and include, but are not limited to CscA from *E. coli* wild-type strain EC3132 (set forth in SEQ ID NO:28), encoded by gene *cscA* (coding sequence set forth in SEQ ID NO:27), CscA from *E. coli* ATCC®13281 (set forth in SEQ ID NO:30),

encoded by gene *cscA* (coding sequence set forth in SEQ ID NO:29); BfrA from *Bifidobacterium lactis* strain DSM 10140^T (set forth in SEQ ID NO:32), encoded by gene *bfrA* (coding sequence set forth in SEQ ID NO:31); Suc2p from *Saccharomyces cerevisiae* (set forth in SEQ ID NO:34), encoded by gene *SUC2* (coding sequence set forth in SEQ ID NO:33); ScrB from *Corynebacterium glutamicum* (set forth in SEQ ID NO:36), encoded by gene *scrB* (coding sequence set forth in SEQ ID NO:35); ScrB from *Pseudomonas fluorescens Pf5* (set forth in SEQ ID NO:134), encoded by gene *scrB* (coding sequence set forth in SEQ ID NO:133), FruP from *Bacillus licheniformis* 14580 (set forth in SEQ ID NO:136), encoded by gene *fruA* (coding sequence set forth in SEQ ID NO:135), sucrose phosphorylase from *Leuconostoc mesenteroides* DSM 20193 (set forth in SEQ ID NO:38), coding sequence of encoding gene set forth in SEQ ID NO:37; and sucrose phosphorylase from *Bifidobacterium adolescentis* DSM 20083 (set forth in SEQ ID NO:40), encoded by gene *sucP* (coding sequence set forth in SEQ ID NO:39).

In some embodiments, the polypeptide having sucrose hydrolase activity is classified as EC 3.2.1.26 or EC 2.4.1.7.

In some embodiments, the polypeptide having sucrose hydrolase activity has at least 95% sequence identity, based on the Clustal W method of alignment, to an amino acid sequence as set forth in SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32, SEQ ID NO:34, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:40, SEQ ID NO:134, or SEQ ID NO:136.

In some embodiments, the polypeptide having sucrose hydrolase activity corresponds substantially to the amino acid sequence set forth in SEQ ID NO:30.

The recombinant bacteria disclosed herein may further comprise in their genome or on at least one recombinant construct, a nucleotide sequence encoding a polypeptide having fructokinase activity to enable the bacteria to utilize the fructose produced by the hydrolysis of sucrose. Polypeptides having fructokinase activity include fructokinases (designated EC 2.7.1.4) and various hexose kinases having fructose

phosphorylating activity (EC 2.7.1.3 and EC 2.7.1.1). Fructose phosphorylating activity may be exhibited by hexokinases and ketohexokinases. Representative genes encoding polypeptides from a variety of microorganisms, which may be used to construct the recombinant bacteria disclosed herein, are listed in Table 2. One skilled in the art will know that proteins that are substantially similar to a protein which is able to phosphorylate fructose (such as encoded by the genes listed in Table 2) may also be used.

Table 2

Sequences Encoding Enzymes with Fructokinase Activity

Source	Gene Name	EC Number	Nucleotide SEQ ID NO:	Protein SEQ ID NO:
Agrobacterium tumefaciens	*scrK* (fructokinase)	2.7.1.4	41	42
Streptococcus mutans	*scrK* (fructokinase)	2.7.1.4	43	44
Escherichia coli	*scrK* (fructokinase	2.7.1.4	45	46
Klebsiella pneumoniae	*scrK* (fructokinase	2.7.1.4	47	48
Escherichia coli	*cscK* (fructokinase)	2.7.1.4	49	50
Enterococcus faecalis	*cscK* (fructokinase)	2.7.1.4	51	52
Saccharomyces cerevisiae	*HXK1* (hexokinase)	2.7.1.1	53	54
Saccharomyces cerevisiae	*HXK2* (hexokinase)	2.7.1.1	55	56

In some embodiments, the polypeptide having fructokinase activity is classified as EC 2.7.1.4, EC 2.7.1.3, or EC 2.7.1.1.

In some embodiments, the polypeptide having fructokinase activity has at least 95% sequence identity, based on the Clustal W method of alignment, to an amino acid sequence as set forth in SEQ ID NO:42, SEQ ID NO:44, SEQ ID NO:46, SEQ ID NO:48, SEQ ID NO:50, SEQ ID NO:52, SEQ ID NO:54, or SEQ ID NO:56.

In some embodiments, the polypeptide having fructokinase activity

has the amino acid sequence set forth in SEQ ID NO:50.

The coding sequence of the genes encoding polypeptides having sucrose transporter activity and polypeptides having sucrose hydrolase activity may be used to isolate nucleotide sequences encoding homologous polypeptides from the same or other microbial species. For example, homologs of the genes may be identified using sequence analysis software, such as BLASTN, to search publically available nucleic acid sequence databases. Additionally, the isolation of homologous genes using sequence-dependent protocols is well known in the art. Examples of sequence-dependent protocols include, but are not limited to, methods of nucleic acid hybridization, and methods of DNA and RNA amplification as exemplified by various uses of nucleic acid amplification technologies (e.g. polymerase chain reaction (PCR), Mullis et al., U.S. Patent No. 4,683,202; ligase chain reaction (LCR), Tabor, S. et al., *Proc. Acad. Sci.* USA 82, 1074, 1985); or strand displacement amplification (SDA), Walker, et al., *Proc. Natl. Acad. Sci. U.S.A.*, 89: 392, (1992)). For example, the nucleotide sequence encoding the polypeptides described above may be employed as a hybridization probe for the identification of homologs.

One of ordinary skill in the art will appreciate that genes encoding these polypeptides isolated from other sources may also be used in the recombinant bacteria disclosed herein. Additionally, variations in the nucleotide sequences encoding the polypeptides may be made without affecting the amino acid sequence of the encoded polypeptide due to codon degeneracy, and that amino acid substitutions, deletions or additions that produce a substantially similar protein may be included in the encoded protein.

The nucleotide sequences encoding the polypeptides having sucrose transporter activity and polypeptides having sucrose hydrolase activity may be isolated using PCR (see, e.g., U.S. Patent No. 4,683,202) with primers designed to bound the desired sequence. Other methods of gene isolation are well known to one skilled in the art such as by using degenerate primers or heterologous probe

hybridization. The nucleotide sequences can also be chemically synthesized or purchased from vendors such as DNA2.0 Inc. (Menlo Park, CA), Integrated DNA Technologies (Coralville, Iowa), and GenScript USA Inc. (Piscataway, NJ). The nucleotide sequences may be codon optimized for expression in the desired host cell.

Expression of the polypeptides may be effected using one of many methods known to one skilled in the art. For example, the nucleotide sequences encoding the polypeptides described above may be introduced into the bacterium on at least one multicopy plasmid, or by integrating one or more copies of the coding sequences into the host genome. The nucleotide sequences encoding the polypeptides may be introduced into the host bacterium separately (e.g., on separate plasmids) or in any combination (e.g., on a single plasmid).

The introduced coding regions that are either on a plasmid(s) or in the genome may be expressed from at least one highly active promoter. An integrated coding region may either be introduced as a part of a chimeric gene having its own promoter, or it may be integrated adjacent to a highly active promoter that is endogenous to the genome or in a highly expressed operon. Suitable promoters include, but are not limited to, *CYC1*, *HIS3*, *GAL1*, *GAL10*, *ADH1*, *PGK*, *PHO5*, *GAPDH*, *ADC1*, *TRP1*, *URA3*, *LEU2*, *ENO*, and *lac*, *ara*, *tet*, *trp*, IP_L, IP_R, *T7*, *tac*, and *trc* (useful for expression in *Escherichia coli*) as well as the *amy*, *apr*, *npr* promoters and various phage promoters useful for expression in *Bacillus*. The promoter may also be the *Streptomyces lividans* glucose isomerase promoter or a variant thereof, described by Payne et al. (U.S. Patent No. 7,132,527).

In some embodiments, the recombinant bacteria disclosed herein are capable of producing glycerol. Biological processes for the preparation of glycerol using carbohydrates or sugars are known in yeasts and in some bacteria, other fungi, and algae. Both bacteria and yeasts produce glycerol by converting glucose or other carbohydrates through the fructose-1,6-bisphosphate pathway in glycolysis. In the

method of producing glycerol disclosed herein, host bacteria may be used that naturally produce glycerol. In addition, bacteria may be engineered for production of glycerol and glycerol derivatives. The capacity for glycerol production from a variety of substrates may be provided through the expression of the enzyme activities glycerol-3-phosphate dehydrogenase (G3PDH) and/or glycerol-3-phosphatase as described in U.S. Patent No. 7,005,291. Genes encoding these proteins that may be used for expressing the enzyme activities in a host bacterium are described in U.S. Patent No. 7,005,291. Suitable examples of genes encoding polypeptides having glycerol-3-phosphate dehydrogenase activity include, but are not limited to, *GPD1* from *Saccharomyces cerevisiae* (coding sequence set forth in SEQ ID NO:1, encoded protein sequence set forth in SEQ ID NO:2) and *GPD2* from *Saccharomyces cerevisiae* (coding sequence set forth in SEQ ID NO:3, encoded protein sequence set forth in SEQ ID NO:4). Suitable examples of genes encoding polypeptides having glycerol-3-phosphatase activity include, but are not limited to, *GPP1* from *Saccharomyces cerevisiae* (coding sequence set forth in SEQ ID NO:5, encoded protein sequence set forth in SEQ ID NO:6) and *GPP2* from *Saccharomyces cerevisiae* (coding sequence set forth in SEQ ID NO:7, encoded protein sequence set forth in SEQ ID NO:8).

Increased production of glycerol may be attained through reducing expression of target endogenous genes. Down-regulation of endogenous genes encoding glycerol kinase and glycerol dehydrogenase activities further enhance glycerol production as described in U.S. Patent No. 7,005,291. Increased channeling of carbon to glycerol may be accomplished by reducing the expression of the endogenous gene encoding glyceraldehyde 3-phosphate dehydrogenase, as described in U.S. Patent No. 7,371,558. Down-regulation may be accomplished by using any method known in the art, for example, the methods described above for down-regulation of genes of the PTS system.

Glycerol provides a substrate for microbial production of useful

products. Examples of such products, i.e., glycerol derivatives include, but are not limited to, 3-hydroxypropionic acid, methylglyoxal, 1,2-propanediol, and 1,3-propanediol.

In some embodiments, the recombinant bacteria disclosed herein are capable of producing 1,3-propanediol. The glycerol derivative 1,3-propanediol is a monomer having potential utility in the production of polyester fibers and the manufacture of polyurethanes and cyclic compounds. 1,3-Propanediol can be produced by a single microorganism by bioconversion of a carbon substrate other than glycerol or dihydroxyacetone, as described in U.S. Patent No. 5,686,276. In this bioconversion, glycerol is produced from the carbon substrate, as described above. Glycerol is converted to the intermediate 3-hydroxypropionaldehyde by a dehydratase enzyme, which can be encoded by the host bacterium or can be introduced into the host by recombination. The dehydratase can be glycerol dehydratase (E.C. 4.2.1.30), diol dehydratase (E.C. 4.2.1.28) or any other enzyme able to catalyze this conversion. A suitable example of genes encoding the "α" (alpha), "β" (beta), and "γ" (gamma) subunits of a glycerol dehydratase include, but are not limited to *dhaB1* (coding sequence set forth in SEQ ID NO:9), *dhaB2* (coding sequence set forth in SEQ ID NO:11), and *dhaB3* (coding sequence set forth in SEQ ID NO:13), respectively, from *Klebsiella pneumoniae*. The further conversion of 3-hydroxypropionaldehyde to 1,3-propandeiol can be catalyzed by 1,3-propanediol dehydrogenase (E.C. 1.1.1.202) or other alcohol dehydrogenases. A suitable example of a gene encoding a 1,3-propanediol dehydrogenase is *dhaT* from *Klebsiella pneumoniae* (coding sequence set forth in SEQ ID NO:57, encoded protein sequence set forth in SEQ ID NO:58).

Bacteria can be recombinantly engineered to provide more efficient production of glycerol and the glycerol derivative 1,3-propanediol. For example, U.S. Patent No. 7,005,291 discloses transformed microorganisms and a method for production of glycerol and 1,3-propanediol with advantages derived from expressing exogenous activities

of one or both of glycerol-3-phosphate dehydrogenase and glycerol-3-phosphate phosphatase while disrupting one or both of endogenous activities glycerol kinase and glycerol dehydrogenase.

U.S. Patent No. 6,013,494 describes a process for the production of 1,3-propanediol using a single microorganism comprising exogenous glycerol-3-phosphate dehydrogenase, glycerol-3-phosphate phosphatase, dehydratase, and 1,3-propanediol oxidoreductase (e.g., *dhaT*). U.S. Patent No. 6,136,576 discloses a method for the production of 1,3-propanediol comprising a recombinant microorganism further comprising a dehydratase and protein X (later identified as being a dehydratase reactivation factor peptide).

U.S. Patent No. 6,514,733 describes an improvement to the process where a significant increase in titer (grams product per liter) is obtained by virtue of a non-specific catalytic activity (distinguished from 1,3-propanediol oxidoreductase encoded by *dhaT*) to convert 3-hydroxypropionaldehyde to 1,3-propanediol. Additionally, U.S. Patent No. 7,132,527 discloses vectors and plasmids useful for the production of 1,3-propanediol.

Increased production of 1,3-propanediol may be achieved by further modifications to a host bacterium, including down-regulating expression of some target genes and up-regulating, expression of other target genes, as described in U.S. Patent No. 7,371,558. For utilization of glucose as a carbon source in a PTS minus host, expression of glucokinase activity may be increased.

Additional genes whose increased or up-regulated expression increases 1,3-propanediol production include genes encoding:

- phosphoenolpyruvate carboxylase typically characterized as EC 4.1.1.31
- cob(I)alamin adenosyltransferase, typically characterized as EC 2.5.1.17
- non-specific catalytic activity that is sufficient to catalyze the interconversion of 3-HPA and 1,3-propanediol, and specifically

excludes 1,3-propanediol oxidoreductase(s), typically these enzymes are alcohol dehydrogenases

Genes whose reduced or down-regulated expression increases 1,3-propanediol production include genes encoding:

- aerobic respiration control protein
- methylglyoxal synthase
- acetate kinase
- phosphotransacetylase
- aldehyde dehydrogenase A
- aldehyde dehydrogenase B
- triosephosphate isomerase
- phosphogluconate dehydratase

In some embodiments, the recombinant bacteria disclosed herein are capable of producing 3-hydroxypropionic acid. 3-Hydroxypropionic acid has utility for specialty synthesis and can be converted to commercially important intermediates by known art in the chemical industry, e.g., acrylic acid by dehydration, malonic acid by oxidation, esters by esterification reactions with alcohols, and 1,3-propanediol by reduction. 3-Hydroxypropionic acid may be produced biologically from a fermentable carbon source by a single microorganism, as described in copending and commonly owned U.S. Patent No. 2011/0144377 A1. In one representative biosynthetic pathway, a carbon substrate is converted to 3-hydroxypropionaldehyde, as described above for the production of 1,3-propanediol. The 3-hydroxypropionaldehyde is converted to 3-hydroxypropionic acid by an aldehyde dehydrogenase. Suitable examples of aldehyde dehydrogenases include, but are not limited to, AldB (SEQ ID NO:16), encoded by the *E. coli* gene *aldB* (coding sequence set forth in SEQ ID NO:15); AldA (SEQ ID NO:18), encoded by the *E. coli* gene *aldA* (coding sequence set forth in SEQ ID NO:17); and AldH (SEQ ID NO:20), encoded by the *E. coli* gene *aldH* (coding sequence as set forth in SEQ ID NO:19).

Many of the modifications described above to improve 1,3-

propanediol production by a recombinant bacterium can also be made to improve 3-hydroxypropionic acid production. For example, the elimination of glycerol kinase prevents glycerol, formed from G3P by the action of G3P phosphatase, from being re-converted to G3P at the expense of ATP. Also, the elimination of glycerol dehydrogenase (for example, *gldA*) prevents glycerol, formed from DHAP by the action of NAD-dependent glycerol-3-phosphate dehydrogenase, from being converted to dihydroxyacetone. Mutations can be directed toward a structural gene so as to impair or improve the activity of an enzymatic activity or can be directed toward a regulatory gene, including promoter regions and ribosome binding sites, so as to modulate the expression level of an enzymatic activity.

Up-regulation or down-regulation may be achieved by a variety of methods which are known to those skilled in the art. It is well understood that up-regulation or down-regulation of a gene refers to an alteration in the level of activity present in a cell that is derived from the protein encoded by that gene relative to a control level of activity, for example, by the activity of the protein encoded by the corresponding (or non-altered) wild-type gene.

Specific genes involved in an enzyme pathway may be up-regulated to increase the activity of their encoded function(s). For example, additional copies of selected genes may be introduced into the host cell on multicopy plasmids such as pBR322. Such genes may also be integrated into the chromosome with appropriate regulatory sequences that result in increased activity of their encoded functions. The target genes may be modified so as to be under the control of non-native promoters or altered native promoters. Endogenous promoters can be altered *in vivo* by mutation, deletion, and/or substitution.

Alternatively, it may be useful to reduce or eliminate the expression of certain genes relative to a given activity level. Methods of down-regulating (disrupting) genes are known to those of skill in the art.

Down-regulation can occur by deletion, insertion, or alteration of

coding regions and/or regulatory (promoter) regions. Specific down regulations may be obtained by random mutation followed by screening or selection, or, where the gene sequence is known, by direct intervention by molecular biology methods known to those skilled in the art. A particularly useful, but not exclusive, method to effect down-regulation is to alter promoter strength.

Furthermore, down-regulation of gene expression may be used to either prevent expression of the protein of interest or result in the expression of a protein that is non-functional. This may be accomplished for example, by 1) deleting coding regions and/or regulatory (promoter) regions, 2) inserting exogenous nucleic acid sequences into coding regions and/regulatory (promoter) regions, and 3) altering coding regions and/or regulatory (promoter) regions (for example, by making DNA base pair changes). Specific disruptions may be obtained by random mutation followed by screening or selection, or, in cases where the gene sequences in known, specific disruptions may be obtained by direct intervention using molecular biology methods know to those skilled in the art. A particularly useful method is the deletion of significant amounts of coding regions and/or regulatory (promoter) regions.

Methods of altering recombinant protein expression are known to those skilled in the art, and are discussed in part in Baneyx, *Curr. Opin. Biotechnol.* (1999) 10:411; Ross, et al., *J. Bacteriol.* (1998) 180:5375; deHaseth, et al., *J. Bacteriol.* (1998) 180:3019; Smolke and Keasling, *Biotechnol. Bioeng.* (2002) 80:762; Swartz, *Curr. Opin. Biotech. (2001)* 12:195; and Ma, et al., *J. Bacteriol.* (2002) 184:5733.

Recombinant bacteria containing the necessary changes in gene expression for metabolizing sucrose in the production of microbial products including glycerol and glycerol derivatives, as described above, may be constructed using techniques well known in the art.

The construction of the recombinant bacteria disclosed herein may be accomplished using a variety of vectors and transformation and expression cassettes suitable for the cloning, transformation and

expression of coding regions that confer the ability to utilize sucrose in the production of glycerol and its derivatives in a suitable host microorganism. Suitable vectors are those which are compatible with the bacterium employed. Suitable vectors can be derived, for example, from a bacterium, a virus (such as bacteriophage T7 or a M-13 derived phage), a cosmid, a yeast or a plant. Protocols for obtaining and using such vectors are known to those skilled in the art (Sambrook *et al.*, *supra*).

Initiation control regions, or promoters, which are useful to drive expression of coding regions for the instant invention in the desired host bacterium are numerous and familiar to those skilled in the art. Virtually any promoter capable of driving expression is suitable for use herein. For example, any of the promoters listed above may be used.

Termination control regions may also be derived from various genes native to the preferred hosts. Optionally, a termination site may be unnecessary; however, it is most preferred if included.

For effective expression of the instant polypeptides, nucleotide sequences encoding the polypeptides are linked operably through initiation codons to selected expression control regions such that expression results in the formation of the appropriate messenger RNA.

Particularly useful are the vectors pSYCO101, pSYCO103, pSYCO106, and pSYCO109, described in U.S. Patent No. 7,371,558, and pSYCO400/AGRO, described in U.S. Patent No. 7,524,660. The essential elements of these vectors are derived from the *dha* regulon isolated from *Klebsiella pneumoniae* and from *Saccharomyces cerevisiae*. Each vector contains the open reading frames *dhaB1*, *dhaB2*, *dhaB3*, *dhaX* (coding sequence set forth in SEQ ID NO:59; encoded polypeptide sequence set forth in SEQ ID NO:60), *orfX*, *DAR1*, and *GPP2* arranged in three separate operons. The nucleotide sequences of pSYCO101, pSYCO103, pSYCO106, pSYCO109, and pSYCO400/AGRO are set forth in SEQ ID NO:62, SEQ ID NO:63, SEQ ID NO:64, SEQ ID NO:65, and SEQ ID NO:66, respectively. The differences between the vectors are illustrated in the chart below [the prefix "p-" indicates a promoter; the open reading

frames contained within each "( )" represent the composition of an operon]:

pSYCO101 (SEQ ID NO:62):

p-trc (Dar1_GPP2) in opposite orientation compared to the other 2 pathway operons,

p-1.6 long GI (dhaB1_dhaB2_dhaB3_dhaX), and

p-1.6 long GI (orfY_orfX_orfW).

pSYCO103 (SEQ ID NO:63):

p-trc (Dar1_GPP2) same orientation compared to the other 2 pathway operons,

p-1.5 long GI (dhaB1_dhaB2_dhaB3_dhaX), and

p-1.5 long GI (orfY_orfX_orfW).

pSYCO106 (SEQ ID NO:64):

p-trc (Dar1_GPP2) same orientation compared to the other 2 pathway operons,

p-1.6 long GI (dhaB1_dhaB2_dhaB3_dhaX), and

p-1.6 long GI (orfY_orfX_orfW).

pSYCO109 (SEQ ID NO:65):

p-trc (Dar1_GPP2) same orientation compared to the other 2 pathway operons,

p-1.6 long GI (dhaB1_dhaB2_dhaB3_dhaX), and

p-1.6 long GI (orfY_orfX).

pSYCO400/AGRO (SEQ ID NO:66):

p-trc (Dar1_GPP2) same orientation compared to the other 2 pathway operons,

p-1.6 long GI (dhaB1_dhaB2_dhaB3_dhaX), and

p-1.6 long GI (orfY_orfX).

p-1.20 short/long GI (scrK) opposite orientation compared to the pathway operons.

Once suitable expression cassettes are constructed, they are used to transform appropriate host bacteria. Introduction of the cassette containing the coding regions into the host bacterium may be

accomplished by known procedures such as by transformation (e.g., using calcium-permeabilized cells, or electroporation) or by transfection using a recombinant phage virus (Sambrook et al., *supra*). Expression cassettes may be maintained on a stable plasmid in a host cell. In addition, expression cassettes may be integrated into the genome of the host bacterium through homologous or random recombination using vectors and methods well known to those skilled in the art. Site-specific recombination systems may also be used for genomic integration of expression cassettes.

In addition to the cells exemplified, cells having single or multiple mutations specifically designed to enhance the production of microbial products including glycerol and/or its derivatives may also be used. Cells that normally divert a carbon feed stock into non-productive pathways, or that exhibit significant catabolite repression may be mutated to avoid these phenotypic deficiencies.

Methods of creating mutants are common and well known in the art. A summary of some methods is presented in U.S. Patent No. 7,371,558. Specific methods for creating mutants using radiation or chemical agents are well documented in the art. See, for example, Thomas D. Brock in Biotechnology: *A Textbook of Industrial Microbiology*, Second Edition (1989) Sinauer Associates, Inc., Sunderland, MA., or Deshpande, Mukund V., *Appl. Biochem. Biotechnol. 36*, 227 (1992).

After mutagenesis has occurred, mutants having the desired phenotype may be selected by a variety of methods. Random screening is most common where the mutagenized cells are selected for the ability to produce the desired product or intermediate. Alternatively, selective isolation of mutants can be performed by growing a mutagenized population on selective media where only resistant colonies can develop. Methods of mutant selection are highly developed and well known in the art of industrial microbiology. See, for example, Brock, *Supra*; DeMancilha *et al., Food Chem. 14*, 313 (1984).

Fermentation media in the present invention comprise sucrose as a

carbon substrate. Other carbon substrates such as glucose and fructose may also be present.

In addition to the carbon substrate, a suitable fermentation medium contains, for example, suitable minerals, salts, cofactors, buffers and other components, known to those skilled in the art, suitable for the growth of the cultures and promotion of the enzymatic pathway necessary for production of glycerol and its derivatives, for example 1,3-propanediol. Particular attention is given to Co(II) salts and/or vitamin B_{12} or precursors thereof in production of 1,3-propanediol.

Adenosyl-cobalamin (coenzyme B_{12}) is an important cofactor for dehydratase activity. Synthesis of coenzyme B_{12} is found in prokaryotes, some of which are able to synthesize the compound *de novo*, for example, *Escherichia blattae*, *Klebsiella* species, *Citrobacter* species, and *Clostridium* species, while others can perform partial reactions. *E. coli*, for example, cannot fabricate the corrin ring structure, but is able to catalyze the conversion of cobinamide to corrinoid and can introduce the 5'-deoxyadenosyl group. Thus, it is known in the art that a coenzyme B_{12} precursor, such as vitamin B_{12}, needs be provided in *E. coli* fermentations. Vitamin B_{12} may be added continuously to *E. coli* fermentations at a constant rate or staged as to coincide with the generation of cell mass, or may be added in single or multiple bolus additions.

Although vitamin B_{12} is added to the transformed *E. coli* described herein, it is contemplated that other bacteria, capable of *de novo* vitamin B_{12} biosynthesis will also be suitable production cells and the addition of vitamin B_{12} to these bacteria will be unnecessary.

Typically bacterial cells are grown at 25 to 40 °C in an appropriate medium containing sucrose. Examples of suitable growth media for use herein are common commercially prepared media such as Luria Bertani (LB) broth, Sabouraud Dextrose (SD) broth or Yeast medium (YM) broth. Other defined or synthetic growth media may also be used, and the

appropriate medium for growth of the particular bacterium will be known by someone skilled in the art of microbiology or fermentation science. The use of agents known to modulate catabolite repression directly or indirectly, e.g., cyclic adenosine 2':3'-monophosphate, may also be incorporated into the reaction media. Similarly, the use of agents known to modulate enzymatic activities (e.g., methyl viologen) that lead to enhancement of 1,3-propanediol production may be used in conjunction with or as an alternative to genetic manipulations with 1,3-propanediol production strains.

Suitable pH ranges for the fermentation are between pH 5.0 to pH 9.0, where pH 6.0 to pH 8.0 is typical as the initial condition.

Reactions may be performed under aerobic, anoxic, or anaerobic conditions depending on the requirements of the recombinant bacterium. Fed-batch fermentations may be performed with carbon feed, for example, carbon substrate, limited or excess.

Batch fermentation is a commonly used method. Classical batch fermentation is a closed system where the composition of the medium is set at the beginning of the fermentation and is not subject to artificial alterations during the fermentation. Thus, at the beginning of the fermentation, the medium is inoculated with the desired bacterium and fermentation is permitted to occur adding nothing to the system. Typically, however, "batch" fermentation is batch with respect to the addition of carbon source, and attempts are often made at controlling factors such as pH and oxygen concentration. In batch systems, the metabolite and biomass compositions of the system change constantly up to the time the fermentation is stopped. Within batch cultures, cells moderate through a static lag phase to a high growth log phase and finally to a stationary phase where growth rate is diminished or halted. If untreated, cells in the stationary phase will eventually die. Cells in log phase generally are responsible for the bulk of production of end product or intermediate.

A variation on the standard batch system is the Fed-Batch system.

Fed-Batch fermentation processes are also suitable for use herein and comprise a typical batch system with the exception that the substrate is added in increments as the fermentation progresses. Fed-Batch systems are useful when catabolite repression is apt to inhibit the metabolism of the cells and where it is desirable to have limited amounts of substrate in the media. Measurement of the actual substrate concentration in Fed-Batch systems is difficult and is therefore estimated on the basis of the changes of measurable factors such as pH, dissolved oxygen and the partial pressure of waste gases such as CO_2. Batch and Fed-Batch fermentations are common and well known in the art and examples may be found in Brock, *supra*.

Continuous fermentation is an open system where a defined fermentation medium is added continuously to a bioreactor and an equal amount of conditioned medium is removed simultaneously for processing. Continuous fermentation generally maintains the cultures at a constant high density where cells are primarily in log phase growth.

Continuous fermentation allows for the modulation of one factor or any number of factors that affect cell growth or end product concentration. For example, one method will maintain a limiting nutrient such as the carbon source or nitrogen level at a fixed rate and allow all other parameters to moderate. In other systems, a number of factors affecting growth can be altered continuously while the cell concentration, measured by the turbidity of the medium, is kept constant. Continuous systems strive to maintain steady state growth conditions, and thus the cell loss due to medium being drawn off must be balanced against the cell growth rate in the fermentation. Methods of modulating nutrients and growth factors for continuous fermentation processes as well as techniques for maximizing the rate of product formation are well known in the art of industrial microbiology and a variety of methods are detailed by Brock, *supra*.

It is contemplated that the present invention may be practiced using batch, fed-batch or continuous processes and that any known mode of

fermentation would be suitable. Additionally, it is contemplated that cells may be immobilized on a substrate as whole cell catalysts and subjected to fermentation conditions for production of glycerol and glycerol derivatives, such as 1,3-propanediol.

In some embodiments, a process for making glycerol, 1,3-propanediol, and/or 3-hydroxypropionic acid from sucrose is provided. The process comprises the steps of culturing a recombinant bacterium, as described above, in the presence of sucrose, and optionally recovering the glycerol, 1,3-propanediol, and/or 3-hydroxypropionic acid produced. The product may be recovered using methods known in the art. For example, solids may be removed from the fermentation medium by centrifugation, filtration, decantation, or the like. Then, the product may be isolated from the fermentation medium, which has been treated to remove solids as described above, using methods such as distillation, liquid-liquid extraction, or membrane-based separation.

EXAMPLES

The present invention is further defined in the following Examples. It should be understood that these Examples, while indicating preferred embodiments of the invention, are given by way of illustration only. From the above discussion and these Examples, one skilled in the art can ascertain the essential characteristics of this invention, and without departing from the spirit and scope thereof, can make various changes and modifications of the invention to adapt it to various uses and conditions.

GENERAL METHODS

Standard recombinant DNA and molecular cloning techniques described in the Examples are well known in the art and are described by Sambrook, J., Fritsch, E. F. and Maniatis, T. *Molecular Cloning: A Laboratory Manual*; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, (1989) (Maniatis) and by T. J. Silhavy, M. L. Bennan, and L. W.

Enquist, *Experiments with Gene Fusions*, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. (1984) and by Ausubel, F. M. *et al.*, *Current Protocols in Molecular Biology*, pub. by Greene Publishing Assoc. and Wiley-Interscience (1987).

The meaning of abbreviations is as follows: "sec" means second(s), "min" means minute(s), "h" means hour(s), "nm" means nanometers, "µL" means microliter(s), "mL" means milliliter(s), "L" means liter(s), "mM" means millimolar, "M" means molar, "g" means gram(s), "µg" means microgram(s), "bp" means base pair(s), "kbp" means kilobase pair(s), "rpm" means revolutions per minute, "ATCC" means American Type Culture Collection, Manassas, VA, "dH_2O" means distilled water.

Media and Culture Conditions:

Materials and methods suitable for the maintenance and growth of bacterial cultures are well known in the art. Techniques suitable for use in the following Examples may be found as set out in *Manual of Methods for General Bacteriology* (Phillipp Gerhardt, R. G. E. Murray, Ralph N. Costilow, Eugene W. Nester, Willis A. Wood, Noel R. Krieg and G. Briggs Phillips, eds), American Society for Microbiology, Washington, DC. (1994)) or by Thomas D. Brock in *Biotechnology: A Textbook of Industrial Microbiology*, Second Edition, Sinauer Associates, Inc., Sunderland, MA (1989). All reagents, restriction enzymes and materials described for the growth and maintenance of bacterial cells may be obtained from Aldrich Chemicals (Milwaukee, WI), BD Diagnostic Systems (Sparks, MD), Life Technologies (Rockville, MD), New England Biolabs (Beverly, MA), or Sigma Chemical Company (St. Louis, MO), unless otherwise specified.

LB (Luria Bertani) medium contains following per liter of medium: Bacto-tryptone (10 g), Bacto-yeast extract (5 g), and NaCl (10 g). Supplements were added as described in the Examples below. All additions were pre-sterilized before they were added to the medium.

Molecular Biology Techniques:

Restriction enzyme digestions, ligations, transformations, and methods for agarose gel electrophoresis were performed as described in Sambrook, J., et al., *Molecular Cloning: A Laboratory Manual*, Second Edition, Cold Spring Harbor Laboratory Press (1989). Polymerase Chain Reactions (PCR) techniques were found in White, B., *PCR Protocols: Current Methods and Applications*, Volume 15 (1993), Humana Press Inc., New York. NY.

EXAMPLES 1 AND 2

Recombinant *E. coli* Strain Comprising a Variant CscB Sucrose Transporter Having a R300A Mutation

The purpose of these Examples was to construct a recombinant *E. coli* strain containing a variant CscB sucrose transport gene (coding sequence set forth in SEQ ID NO:99), encoding an R300A variant of CscB (SEQ ID NO:100), and to demonstrate sucrose transport by facilitated diffusion. The protein encoded by the mutant sucrose transport gene was altered in a residue required for H^+ translocation, thus eliminating H^+/sucrose symport (i.e., active transport of sucrose).

Construction of Expression Vectors:

Two expression vectors were constructed, one using promoter element P1.20 and the second using promoter element P1.5. P1.20 and P1.5 refer to promoter elements derived from the *Streptomyces lividans* glucose isomerase promoter as described in U.S. Patent No. 7,132,527. These two promoters differ from each other by one base in the -35 region such that P1.20 confers lower expression than does P1.5.

The promoter/multiple cloning site/double terminator regions were synthesized by Integrated DNA Technologies (Coralville, Iowa) and cloned into their pIDTsmart vector, resulting in the construction of plasmids named pDMWP1 and pDMWP3. The seqeuces of the synthesized regions for vectors pDMWP1 and pDMWP3 are set forth in SEQ ID

NO:113 and SEQ ID NO:114, respectively.

A plasmid referred to herein as pDMWP4 was used as the backbone for subsequent constructs. Plasmid pDMWP4 was constructed from plasmid pBR322 by modifying restriction sites as follows. A Sca1 site and a KpnI site on the 5' end of the TetR gene and an additional KpnI site at the 3' end of the TetR gene were introduced into plasmid pBR322. Additionally, a KpnI site was removed from the middle of the AmpR gene as well. All sites were either added or removed using Stratagene's QuikChange® kits (Stragene, La Jolla, CA) following manufacturer's protocols.

Plasmids pDMWP1 and DMWP3 were digested with EcoRI and KpnI. The resulting 438 bp fragment from each construct was individually cloned into pDMWP4, also digested with EcoRI and KpnI, to complete plasmids pDMWP10 and pDMWP12, which are also referred to herein as pBR*P1.5 and pBR*1.20, respectively.

The R300A variant of CscB was given the allele name, *cscB3*. This mutation was introduced into plasmid pBHRcscBKA (described in U.S. Patent Application Publication No. 2011/0136190, Example 1) by site-directed mutagenesis using Stratagene's QuikChange® Site-Directed Mutagenesis kit following the manufacturer's protocol. Primers ODMWP23 (SEQ ID NO:115) and ODMWP24 (SEQ ID NO:116) were used with plasmid pBHRcscBKA as template in the reaction, creating plasmid pDMWP5. The *cscB3* gene was subsequently amplified from pDMWP5 using primers ODMWP31 (SEQ ID NO:117) and ODMWP32 (SEQ ID NO:118) to add HindIII/ClaI sites. The resulting product was cloned into pBADtopo (Invitrogen, Carlsbad, CA) creating plasmid pDMWP26.

The HindIII/Cla fragment from pDMWP26 was cloned into HindIII/ClaI digested pDMWP12, creating plasmid pDMWP32, which contained promoter P1.20.

The HindIII/Pac fragment from pDMWP32 was cloned into HindIII/Pac digested pDMWP10, creating pDMWP73, which contained

promoter P1.5.

Construction of *E. coli* strains with or without Expression of *cscB3*:

E. coli strain PDO3513, an E. coli K12 strain [FM5 yihP:cscA+K+B-(Δ61-353, kanR)] that does not have sucrose transporter function, but possesses genes encoding sucrose invertase and fructokinase for downstream metabolism was used as the host strain. *E. coli* strain PDO3513 was constructed from an *E. coli* strain (referred to herein as PDO3085) containing the wild type *cscAKB* gene cluster from *E. coli* ATCC®13281, integrated at the *yihP* gene in *E. coli* strain FM5 (ATCC® No. 53911). The *cscAKB* gene cluster (SEQ ID NO:61) was integrated at the *yihP* location in *E. coli* strain FM5 (ATCC® No. 53911) by the Lambda Red method. The *cscAKB* gene cluster was amplified from plasmid pBHR-cscBKA (SEQ ID NO:119), which was constructed as described in Example 1 of U.S. Patent Application Publication No. 2011/0136190 A1, using yihP cscA primer (SEQ ID NO:120) and yihP cscB primer (SEQ ID NO:121) containing flanking sequences for the *yihP* gene. Plasmid pBHR-cscBKA, linearized by PstI digest, was used as the PCR template. High fidelity PfuUltra ® II Fusion HS DNA polymerase (Stratagene; La Jolla, CA) was used in the PCR reaction. PCR was performed using the following cycling conditions: 95 °C for 2 min; 35 cycles of 95 °C for 30 sec, 60 °C for 30 sec, and 72 °C for 4 min; and then 72 °C for 7 min. The resulting PCR product was stored at 4 °C. The PCR product was purified using a QIAquick PCR Purification kit (Qiagen, Valencia, CA). The purified PCR product was electroporated into *E. coli* strain FM5 containing the pKD46 plasmid (Red recombinase plasmid, GenBank Acc. No. AY048746), encoding lambda recombinases, following the lambda red recombination procedure (Datsenko, K.A. and Wanner, B.L., 2000, *Proc. Natl. Acad. Sci. USA* **97**, 6640-6645). The transformation mixture was plated on MOPS minimal plates containing 10 g/L sucrose. The MOPS minimal plates contained 1XMOPS buffer (Technova, Hollister, CA), 1.32 mM KH_2PO_4 (Technova), 50 µg/L uracil and 1.5 g/L Bacto agar. Plates

were incubated at 37 °C for 2-3 days. Colonies grown on minimal sucrose plates were picked to give *E. coli* strain PDO3085.

The *cscB* gene in the cluster in PDO3085 was then partially deleted by replacing it with a kanamycin resistance cassette. The kanamycin resistance cassette was amplified from the pKD4 template plasmid (Datsenko and Wanner, *Proc. Natl. Acad. Sci. USA* 97:6640-6645, 2000) using cscB61 up kan primer (SEQ ID NO:91) and cscB353 down kan primer (SEQ ID NO:92). High fidelity PfuUltra ® II Fusion HS DNA polymerase (Stratagene; La Jolla, CA) was used in the PCR reaction. PCR was performed using the following cycling conditions: 95 °C for 2 min; 30 cycles of 95 °C for 20 sec, 60 °C for 20 sec, and 72 °C for 1.5 min; and then 72 °C for 3 min. The resulting PCR product was stored at 4 °C. The PCR product was purified using the QIAquick PCR Purification kit (Qiagen). The purified PCR product was electroporated into the PDO3085 strain containing the pKD46 plasmid encoding lambda recombinases following the lambda red recombination procedure. The transformation mixture was plated on LB plates containing 25 µg/mL kanamycin. The kanamycin resistance colonies were checked on MOPS+10 g/L sucrose plates to make sure that they were unable to grow on sucrose. Insertion of the kanamycin resistance cassette between residue 61 and 353 of CscB was confirmed by PCR using cscB 5' primer (SEQ ID NO:93) and cscB 3' primer (SEQ ID NO: 94). The resulting FM5 yihP:cscA+K+B-(Δ61-353, kanR) strain was designated as PDO3513.

Plasmids pDMWP10 (the vector alone) and pDMWP73, carrying the mutant *cscB3* gene, were introduced independently into *E. coli* strain PDO3513. The resultant strains were named PDO2768 and PDO2770, respectively.

Growth Characterization of *E. coli* Strains with or without Expression of *cscB3*:

E. coli strains PDO2768 (Example 2, Comparative) and PDO2770

(Example 1) were grown overnight in LB (Luria Bertani) medium containing 100 µg/mL ampicillin at 37 °C. The next day, these cultures were diluted 1:50 in MOPS minimal medium (Teknova, Half Moon Bay, CA) containing 2 g/L sucrose and 25 µg/mL ampicillin, These cultures were grown at 37 °C with shaking at 250 rpm for 4 hours. The log-phase cultures were diluted 1:100 in the wells of a Bioscreen-C plate (instrument and plates purchased from Growth Curves USA, Piscataway NJ) with150 µL of MOPS minimal medium (Teknova, Half Moon Bay, CA) containing 2 g/L glucose or 16 g/L sucrose. The cultures were grown at 37 °C in triplicate with continuous shaking and the optical density was monitored. The optical densities of the two cultures at 40 hours after inoculation are given in Table 3.

Table 3

Optical Density of Cultures Growing on Glucose or Sucrose at 40 Hours

Carbon Source	Example 1 PDO2770	Example 2, Comparative PDO2768
2 g/L glucose	0.583 ± 0.045	0.572 ± 0.022
16 g/L sucrose	0.892 ± 0.023	0.012 ± 0.003

As can be seen from the data in Table 3, both strains grew well with glucose as a sole carbon source. In contrast, the control strain (i.e., vector only strain) PDO2768 (Example 2, Comparative) was unable to grow with sucrose as sole carbon source, while PDO2770 (Example 1), the strain expressing the mutant *cscB3* gene encoding a sucrose transporter unable to translocate H^+ ion was able to grow with sucrose as sole carbon source. Thus, net translocation of sucrose across the membrane must have occurred without translocation of a H^+ ion.

EXAMPLE 3

Recombinant *E. coli* Strain Comprising a Variant CscB Sucrose

Transporter Having R300A and Q353H Mutations

The purpose of this Example was to construct a recombinant *E. coli* strain containing a variant CscB sucrose transport gene (coding sequence set forth in SEQ ID NO:103), encoding CscB having R300A and Q353H mutations (SEQ ID NO:104), and to demonstrate improved sucrose transport by facilitated diffusion with the additonal mutation conferring a Q353H alteration in *cscB3*.

Construction of Expression Vectors:

For ease of cloning into a smaller vector, the KanR gene from pBHRcscBKAmutB (described in U.S. Patent Application Publication No. 2011/0136190, Example 1) was removed by digesting the plasmid with PstI and religating, creating plasmid pDMWP6. The new vector was 1240 bp smaller than the parent. The mutant *cscB* gene in this vector confers the Q353H variation with improved sucrose transport (Jahreis et al., *J. Bacteriol.* 184:5307-5316, 2002) as compared to the wild type sucrose symporter. It was not known if this mutation would improve sucrose transport by facilitated diffusion.

A mutation conferring the R300A variation was introduced into plasmid pDMWP6 by site-directed mutagenesis using Stratagene's QuikChange® Site-Directed Mutagenesis kit following the manufacturer's protocol. Primers ODMWP23 (SEQ ID NO:115) and ODMWP24 (SEQ ID NO:116) were used with plasmid pDMWP6 as template in the reaction, creating plasmid pDMWP15. The *cscB5* gene (containing R300A and Q353H mutations) was subsequently amplified from pDMWP6 using primers ODMWP31 (SEQ ID NO:117) and ODMWP32 (SEQ ID NO:118) to add HindIII/ClaI sites. The resulting product was cloned into pBADtopo (Invitrogen, Carlsbad, CA), creating plasmid pDMWP27.

The HindIII/Cla fragment from pDMWP27 was cloned into HindIII/ClaI digested pDMWP12, creating pDMWP33, which contained the P1.20 promoter.

The HindIII/Pac fragment from pDMWP33 was cloned into HindIII/Pac digested pDMWP10, creating pDMWP66, which contained the P1.5 promoter.

Construction of *E. coli* Strain with Expression of *cscB5*:

Plasmid pDMWP66 (pBR*p1.5*csc5*) was transformed into strain PDO3513, to give strain PDO2771.

Growth Characterization of *E. coli* Strains with Expression of *cscB3* or *cscB5*:

E. coli strains PDO2770 (with pBR*p1.5*csc3,* described in Examples 1 and 2) and PDO2771 (with pBR*p1.5*csc5)* were grown overnight in LB (Luria Bertani) medium containing 100 µg/mL ampicillin at 37 °C. The next day, these cultures were diluted 1:50 in MOPS minimal medium (Teknova, Half Moon Bay, CA) containing 2 g/L sucrose and 25 µg/mL ampicillin, These cultures were grown at 37 °C with shaking at 250 rpm for 4 hours. The log-phase cultures were diluted 1:100 in the wells of a Bioscreen-C plate (instrument and plates purchased from Growth Curves USA, Piscataway NJ) with 150 µL MOPS minimal medium (Teknova, Half Moon Bay, CA) containing 2 g/L glucose or 16 g/L sucrose. The cultures were grown at 37 °C in triplicate with continuous shaking and the optical density was monitored. The growth on sucrose was much faster in cultures of strain PDO2771 than cultures of strain PDO2770. At 14 hours after inoculation, the optical density of the PDO2770 culture growing on 16 g/L sucrose was 0.060 ± 0.024 while that of the PDO2771 culture growing on 16 g/L sucrose was 0.647 ± 0.009. As a measure of the health of the inoculum cultures, the growth on glucose was measured. Both strains grew well with glucose as a sole carbon source. At 14 hours after inoculation, the optical density of the PDO2770 culture growing on glucose was 0.639 ± 0.037 and the optical density of the PDO2771 culture growing on glucose was 0.693 ± 0.070. These results demonstrate that the strain expressing CscB5, the sucrose transporter with both Q353H and

R300A mutations, was able to grow much better with sucrose as sole carbon source than did the strain expressing CscB3 (R300A) alone. Because the CscB5 protein still carries a mutation in a residue essential for H^+ translocation, it must be transporting sucrose without translocation of a H^+ ion. Thus, the transporter encoded by the gene with the double mutation is an improved facilitated diffusion sucrose transporter.

EXAMPLES 4-6

Growth on Sucrose of Recombinant *E. coli* Strains Comprising Mutant or Wild Type Sucrose Transporters

The purpose of these Examples was to show that a recombinant *E. coli* strain comprising a variant of CscB having R300A and Q353H mutations was able to grow at a wider range of sucrose concentrations than *E. coli* strains comprising the wild type sucrose transporter.

Construction of Expression Vectors:

The wild type *E. coli cscB* gene was originally amplified from pBHRcscBKA (SEQ ID NO:119), described in U.S. Patent Application Publication No. 2011/0136190, Example 1) with primers ODMWP31 (SEQ ID NO:117) and ODMWP32 (SEQ ID NO:118), allowing the addition of both HindIII and ClaI sites at the 5' and 3' ends of the gene, respectively. The PCR fragment was cloned into pBADtopo (Invitrogen, Carlsbad, CA), creating plasmid pDMWP25.

The HindIII/ Cla fragment from pDMWP25 was cloned into HindIII/ Cla digested pDMWP12, creating pDMWP31, which contained promoter P1.20.

The HindIII/Pac digested fragment from pDMWP31 was cloned into the HindIII/Pac digested pDMWP10, creating pDMWP71, which contained promoter P1.5.

Construction of *E. coli* Strains Comprising the Wild Type Sucrose Transporter CscB:

Plasmids pDMWP31 (pBR*p1.20*cscB*) and pDMWP71 (pBR*p1.5*cscB*) were transformed independently into strain PDO3513, to give strains PDO2625 and PDO2769, respectively.

Growth Characterization of Strains with Expression of *cscB5* or Wild Type *cscB*:

The two *E. coli* strains with plasmids encoding the wild type sucrose symporter, PDO2625 (Example 5, Comparative) and PDO2769 (Example 6, Comparative), and a strain with a plasmid carrying the improved sucrose uniporter, PDO2771 (Example 4, with pBR*p1.5*csc5,* described in Example 3), were grown overnight in LB (Luria Bertani) medium containing 100 µg/mL ampicillin at 37 °C. The next day, these cultures were diluted 1:50 in MOPS minimal medium (Teknova, Half Moon Bay, CA) containing 2 g/L sucrose and 25 µg/mL ampicillin, These cultures were grown at 37 °C with shaking at 250 rpm for 4 hours. The log-phase cultures were diluted 1:100 in the wells of a Bioscreen-C plate (instrument and plates purchased from Growth Curves USA, Piscataway NJ) with 150 µL of MOPS minimal medium (Teknova, Half Moon Bay, CA) containing 2 g/L glucose or 2, 4, 8, 16, or 32 g/L sucrose. The cultures were grown at 37 °C in triplicate with continuous shaking and the optical density was monitored. The growth on various concentrations of sucrose was followed. Table 4 shows the optical density at 14 hours after inoculation for the cultures of PDO2771, PDO2625, and PDO2769.

Table 4

Optical Density of Cultures Growing on Glucose or Various Concentrations of Sucrose at 14 Hours

Carbon Source	Example 4 PDO2771 (pBR1.5*cscB5*)	Example 5, Comparative PDO2625 (pBR*p1.20*cscB)*	Example 6, Comparative PDO2769 (pBR*p1.5*cscB)*
2 g/L glucose	0.693 ± 0.070	0.593 ± 0.005	0.668 ± 0.013
2 g/L sucrose	0.071 ± 0.036	0.452 ± 0.028	0.654 ± 0.007
4 g/L sucrose	0.343 ± 0.184	0.719 ± 0.007	0.716 ± 0.011
8 g/L sucrose	0.745 ± 0.004	0.655 ± 0.010	0.094 ± 0.012
16 g/L sucrose	0.647± 0.009	0.107 ± 0.007	0.060 ± 0.005
32 g/L sucrose	0.576 ± 0.006	0.049 ± 0.002	0.058 ± 0.006

As shown by the results in Table 4, all three strains grew well on 2 g/L glucose, indicating that the inoculum cultures were viable. The growth of the PDO2625 strain (Example 5, Comparative) and PDO2769 strain (Example 6, Comparative) was better than that of PDO2771 strain (Example 4) at low sucrose concentrations of 2 or 4 g/L. However at the higher sucrose concentrations of 16 or 32 g/L, PDO2771 maintained good growth while the growth of PDO2525 and PDO2769 was severely inhibited. These results demonstrate that the strain expressing CscB5, the sucrose facilitated diffusion transporter, was able to grow at a wider range of sucrose concentrations than the strains expressing the wild type sucrose symporter. Thus, facilitated diffusion, or uniport, has an advantage of allowing growth under conditions at which the symporter does not allow growth.

EXAMPLES 7-10

PDO Production From Sucrose with a Strain Comprising a Variant of CscB Sucrose Transporter Having R300A and Q353H Mutations and a Strain Comprising the Wild Type Sucrose Transporter CscB

The purpose of these Examples was to show that a recombinant *E. coli* strain comprising a variant of CscB having R300A and Q353H mutations gave better PDO production when grown on sucrose than a recombinant *E. coli* strain comprising the wild type sucrose transporter CscB.

A strain for testing the function of sucrose transporters for PDO production was constructed using PDO producing strain TTab pSYCO400/AGRO. *E. coli* strain TTab pSYCO400/AGRO, a PTS minus strain, was constructed as follows. Strain TTab was generated by deletion of the *aldB* gene from strain TT aldA, described in U.S. Patent No. 7,371,558 (Example 17). Briefly, an *aldB* deletion was made by first replacing 1.5 kbp of the coding region of *aldB* in *E. coli* strain MG1655 with the FRT-CmR-FRT cassette of the pKD3 plasmid (Datsenko and Wanner, *Proc. Natl. Acad. Sci. USA* 97:6640-6645, 2000). A replacement cassette was amplified with the primer pair SEQ ID NO:99 and SEQ ID NO:100 using pKD3 as the template. The primer SEQ ID NO:99 contains 80 bp of homology to the 5'-end of *aldB* and 20 bp of homology to pKD3. Primer SEQ ID NO:100 contains 80 bp of homology to the 3' end of *aldB* and 20 bp homology to pKD3. The PCR products were gel-purified and electroporated into MG1655/pKD46 competent cells (U.S. Patent No. 7,371,558). Recombinant strains were selected on LB (Luria Bertani) plates with 12.5 mg/L of chloramphenicol. The deletion of the *aldB* gene was confirmed by PCR, using the primer pair SEQ ID NO:101 and SEQ ID NO:102. The wild-type strain gave a 1.5 kbp PCR product while the recombinant strain gave a characteristic 1.1 kbp PCR product. A P1 lysate was prepared and used to move the mutation to the TT aldA strain to form the TT aldAΔaldB::Cm strain. A chloramphenicol-resistant clone

was checked by genomic PCR with the primer pair SEQ ID NO:101 and SEQ ID NO:102 to ensure that the mutation was present. The chloramphenicol resistance marker was removed using the FLP recombinase (Datsenko and Wanner, *supra*) to create TTab. Strain TTab was then transformed with pSYCO400/AGRO (set forth in SEQ ID NO:84), described in U.S. Patent No. 7,524,660 (Example 4), to generate strain TTab pSYCO400/AGRO.

As described in the cited references, strain TTab is a derivative of *E. coli* strain FM5 (ATCC® No. 53911) containing the following modifications:

deletion of *glpK, gldA, ptsHI, crr, edd, arcA, mgsA, qor, ackA, pta, aldA* and *aldB* genes;

upregulation of *galP, glk, btuR, ppc,* and *yqhD* genes; and

downregulation of *gapA* gene.

Plasmid pSYCO400/AGRO contains genes encoding a glycerol production pathway (*DAR1* and *GPP2)* and genes encoding a glycerol dehydratase and associated reactivating factor (*dhaB123, dhaX, orfX, orfY)*, as well as a gene encoding a fructokinase (*scrK*).

Strain TTab pSYCO400/AGRO was used as a recipient for P1 transduction. The donor strain was PDO3513, constructed as described in Examples 1 and 2, and selection for growth was on LB plates with 25 µg/mL kanamycin. A colony resistant to kanamycin and spectinomycin was purified and named PDO2737 [TTab/pSYCO400AGRO yihP::cscKBΔ(61-353)KanR&A].

Strain PDO2737 was transformed with plasmids encoding the wild type sucrose transporter, pDMWP31 (pBR*p1.20*cscB*) and pDMWP71 (pBR*p1.5*cscB*) described in Examples 4-6, to yield strains PDO2815 and PDO2818, respectively. In addition, strain PDO2737 was transformed with plasmids encoding a facilitated diffusion sucrose transporter, pDMWP33 (pBR*p1.20*csc5*) and pDMWP66 (pBR*p1.5*csc5*), described in Example 3, to yield strains PDO2965 and PDO2966, respectively.

To test for production of PDO and glycerol, these four *E. coli* strains were grown overnight in L-Broth, Miller's Modification (Teknova, Half Moon Bay, CA) supplemented with 100 mg/L spectinomycin and 100 mg/L ampicillin at 33 °C. These cultures were used to inoculate shake flasks at an optical density of 0.01 units measured at 550 nm in MOPS minimal medium (Teknova, Half Moon Bay, CA) supplemented with 10 g/L sucrose. Vitamin B12 was added to the medium to a concentration of 0.1 mg/L. The cultures were incubated at 34 °C with shaking (225 rpm) for 44 hours. Samples of the cultures were then filtered and used for the determination of the concentrations of sucrose, glycerol and 1,3-propanediol (PDO) in the broth by high performance liquid chromatography.

Chromatographic separation was achieved using an Aminex HPX-87P column (Bio-Rad, Hercules, CA) with an isocratic mobile phase of distilled-deionized water at a flow rate of 0.5 mL/min and a column temperature of 85 °C. Eluted compounds were quantified by refractive index detection with reference to a standard curve prepared from commercially purchased pure compounds dissolved to known concentrations in MOPS minimal medium. Retention times were sucrose at 12.2 min, 1,3-propanediol at 17.9 min, and glycerol at 23.6 min. Table 5 shows the residual sucrose and molar yield of PDO and glycerol (mol PDO + mol glycerol/mol glucose equivalent), in the cultures of these four strains.

Table 5

Sucrose Utilization and PDO and Glycerol Production

Example	Strain	44 hour sucrose g/L	Molar Yield (mol PDO + glycerol/mol glucose equivalent)
Example 7, Comparative	PDO2815 (P1.20cscB)	6.65	0.780
Example 8, Comparative	PDO2818 (P1.5cscB)	9.23	0.660
Example 9	PDO2965 (P1.20cscB5)	4.80	1.014
Example 10	PDO2966 (P1.5cscB5)	1.78	1.066

As can be seen from the results in Table 5, there was more sucrose remaining in the cultures expressing the wild type sucrose transporter CscB (Comparative Examples 7 and 8) than was left in the cultures expressing the facilitated diffusion transporter, CscB5 (Examples 9 and 10), indicating faster sucrose utilization with the facilitated diffusion transporter under these conditions. The molar yield of PDO and glycerol from sucrose was substantially higher for the strains expressing the facilitated diffusion transporter. Thus sucrose transport by facilitated diffusion was shown to be better than with the wild type transporter for PDO and glycerol production.

EXAMPLES 11-14

Recombinant *E. coli* Strain Comprising a Variant CscB Sucrose Transporter Having R300A, Q353H and L61P Mutations

The purpose of these Examples was to demonstrate that recombinant *E. coli* strains comprising variants of CscB having an L61P mutation in addition to an R300A and/or Q353H mutations (SEQ ID NO:106, encoded by SEQ ID NO:105) have improved sucrose transport

by facilitated diffusion. The L61P variation confers improved sucrose transport to the CscB sucrose symporter, as described in copending and commonly owned U.S. Patent Application No. 13/ 210488, but it was not known if this mutation would improve transport by facilitated diffusion, or if the combination of L61P and Q535H would have still further improved transport.

Construction of Expression Vectors:

The *cscB16* allele contains two mutations, L61P and R300A. Plasmid pDMWP32 (described in Examples 1 and 2), which contains the R300A mutation was further mutated to introduce an L61P mutation. The mutation was introduced into pDMWP32 by site directed mutagenesis using Stratagene's QuikChange® Site-Directed Mutagenesis kit, and oligonucleotides ODMWP33 (SEQ ID NO:122) and ODMWP34 (SEQ ID NO:123) following the manufacturer's protocol, creating plasmid pDMWP54.

The *cscB17* allele contains three mutations, L61P, R300A and Q353H. Plasmid pDMWP33 (described in Example 3), which contains the R300A and Q353H mutations was further mutated to introduce an L61P mutation. The mutation was introduced into pDMWP33 by site directed mutagenesis using Stratagene's QuikChange® Site-Directed Mutagenesis kit, and oligonucleotides ODMWP33 (SEQ ID NO:122) and ODMWP34 (SEQ ID NO:123) following the manufacturer's protocol, creating plasmid pDMWP55. The HindIII/Pac fragment from pDMWP55 was cloned into HindIII/Pac digested pDMWP10, to create the P1.5-containing version of the construct, plasmid pDMWP79.

Construction of *E. coli* Strains

Two of the plasmids described above, pDMWP54 (pBR*p1.20*cscB16*) and pDMWP55 (pBR*p1.20*cscB17*), were transformed independently into strain PDO3513, to give strains PDO2636 and PDO2637, respectively. In addition, plasmids pDMWP32

(pBR*1.20cscB3, described in Examples 1 and 2) and pDMWP33 (pBR*1.20cscB5, described in Example 3) were transformed into strain PDO3513, to give strains PDO2626 and PDO2627, respectively.

Growth Characterization of Strains with Expression of *cscB3, cscB5, cscB16, or cscB17*:

The four *E. coli* strains described above were grown overnight in LB (Luria Bertani) medium containing 100 µg/mL of ampicillin at 37 °C. The next day, these cultures were diluted 1:100 in LB medium containing 100 µg/mL of ampicillin, These cultures were grown at 37 °C with shaking at 250 rpm for 4 hours. The log-phase cultures were diluted 1:100 in the wells of a Bioscreen-C plate (instrument and plates purchased from Growth Curves USA, Piscataway NJ) with 150 µL of MOPS minimal medium (Teknova, Half Moon Bay, CA) containing 2 g/L glucose or 2 g/L sucrose. The cultures were grown at 37 °C in triplicate with continuous shaking and the optical density was monitored. Table 6 shows the mean and standard deviation of the optical density readings at 10 hours after inoculation.

Table 6

Growth in Glucose or Sucrose of Strains Expressing Various Sucrose Uniporters Measured by Optical Density at 10 hours.

Example	Strain	cscB allele and variant amino acids	2 g/L glucose	2 g/L sucrose
Example 11	PDO2626	cscB3 (R300A)	0.710 ± 0.007	0.064 ± 0.006
Example 12	PDO2627	cscB5 (R300A Q353H)	0.702 ± 0.005	0.211 ± 0.006
Example 13	PDO2636	cscB16 (L61P R300A)	0.710 ± 0.005	0.609 ± 0.017
Example 14	PDO2637	cscB17 (L61P R300A Q353H)	0.703 ± 0.001	0.732 ± 0.009

As can be seen from the data in Table 6, all four strains grew well on glucose indicating that the inoculum cultures were healthy. Under these growth conditions, there was very little growth of strain PDO2626 expressing the facilitated diffusion transporter CscB3 with the R300A mutation. Comparatively, the growth was dramatically improved in strain PDO2636 expressing CscB16 (L61P and R300A). Likewise, L61P added to R300A Q353H improved growth as seen by comparing the growth of PDO2637 with PDO2627. These results demonstrate that the strain expressing variant sucrose facilitated diffusion transporter with the L61P mutation improved growth with sucrose as sole carbon source. Each of the variant CscB proteins carries the R300A mutation in a residue essential for H^+ translocation, thus each must be transporting sucrose without translocation of a H^+ ion. Accordingly, the transporters encoded

by the genes conferring the L61P variation are improved facilitated diffusion sucrose transporters.

EXAMPLES 15-17

Growth on Sucrose of Recombinant *E. coli* Strains Comprising Mutant or Wild Type Sucrose Transporters

The purpose of these Examples was to show that a recombinant *E. coli* strain comprising a variant of CscB having R300A, Q353H, and L61P mutations was able to grow at a wider range of sucrose concentrations than *E. coli* strains comprising the wild type sucrose transporter.

Construction of *E. coli* Strain Comprising the Variant of CscB Having R300A, Q353H, and L61P Mutations:

Plasmid pDMWP79 (pBR*p1.5*cscB17*, described in Examples 11-14) was transformed into strain PDO3513, to give strain PDO2773.

Growth Characterization of *E. coli* Strains:

Strain PDO2773 (Example 15) and two *E. coli* strains with plasmids encoding the wild type sucrose symporter, PDO2625 (Example 16, Comparative) and PDO2769 (Example 17, Comparative), both described in Examples 4-6, were grown overnight in LB (Luria Bertani) medium containing 100 µg/mL ampicillin at 37 °C. The next day, these cultures were diluted 1:50 in MOPS minimal medium (Teknova, Half Moon Bay, CA) containing 2 g/L sucrose and 25 µg/mL ampicillin, These cultures were grown at 37 °C with shaking at 250 rpm for 4 hours. The log-phase cultures were diluted 1:100 in the wells of a Bioscreen-C plate (instrument and plates purchased from Growth Curves USA, Piscataway NJ) with 150 µL MOPS minimal medium (Teknova, Half Moon Bay, CA) containing 2 g/L glucose or 2, 4, 8, 16, or 32 g/L sucrose. The cultures were grown at 37 °C in triplicate with continuous shaking and the optical density was monitored. The growth on various concentrations of sucrose was

followed. Table 7 shows the optical density at 14 hours after inoculation for the cultures of PDO2773, PDO2625, and PDO2769.

Table 7

Optical Density of Strains Growing on Glucose or Various Concentrations of Sucrose at 14 Hours

	Example 15 PDO2773 (pBR1.5*cscB17*)	Example 16, Comparative PDO2625 pBR*p1.20*cscB*	Example 17, Comparative PDO2769 pBR*p1.5*cscB*
2 g/L glucose	0.669 ± 0.042	0.593 ± 0.005	0.668 ± 0.013
2 g/L sucrose	0.474 ± 0.115	0.452 ± 0.028	0.654 ± 0.007
4 g/L sucrose	0.744 ± 0.043	0.719 ± 0.007	0.716 ± 0.011
8 g/L sucrose	0.657 ± 0.052	0.655 ± 0.010	0.094 ± 0.012
16 g/L sucrose	0.666 ± 0.008	0.107 ± 0.007	0.060 ± 0.005
32 g/L sucrose	0.538 ± 0.015	0.049 ± 0.002	0.058 ± 0.006

As can be seen from the data in Table 7, all three strains grew well on 2 g/L glucose, indicating that the inoculum cultures were viable. The growth of the PDO2625 strain (Example 16, Comparative) and PDO2769 strain (Example 17, Comparative) was similar to the growth of PDO2771 (Example 15) at low sucrose concentrations of 2 or 4 g/L. However at the higher sucrose concentrations of 16 or 32 g/L, PDO2771 maintained good growth while the growth of PDO2525 and PDO2769 was severely inhibited. These results demonstrate that the strain expressing CscB17, the improved sucrose facilitated diffusion transporter with three altered residues, L61P, R300A, Q353H, was able to grow at a much wider range of sucrose concentrations than the strains expressing the wild type sucrose symporter. Thus, this improved facilitated diffusion transporter has an advantage over sucrose transport by a symport mechanism.

EXAMPLES 18-21

PDO Production From Sucrose with a Strain Comprising a Variant CscB Sucrose Transporter Having R300A, Q353H, and L61P Mutations and a Strain Comprising the Wild Type Sucrose Transporter CscB

The purpose of these Examples was to show that a recombinant *E. coli* strain comprising a variant of CscB having R300A, Q353H, and L61P mutations gave better PDO production when grown on sucrose than a recombinant *E. coli* strain comprising the wild type sucrose transporter CscB.

E. coli strain PDO2737 [TTab/pSYCO400AGRO yihP::cscKBΔ(61-353)KanR&A, described in Examples 7-10, was transformed independently with plasmids pDMWP55 (pBR*P1.20cscB17, described in Examples 11-14) and pDMWP79 (pBR*P1.5csscB17, described in Examples 11-14) to make strains PDO2816 and PDO2819, respectively. These two strains and two strains expressing the wild type cscB symporter, PDO2815 and PDO2818 (described in Examples 7-10) were grown overnight in L-Broth, Miller's Modification (Teknova, Half Moon Bay, CA) supplemented with 100 mg/L spectinomycin and 100 mg/L ampicillin at 33 °C. These cultures were used to inoculate shake flasks at an optical density of 0.01 units measured at 550 nm in MOPS minimal medium (Teknova, Half Moon Bay, CA) supplemented with 10 g/L sucrose. Vitamin B12 was added to the medium to a concentration of 0.1 mg/L. The cultures were incubated at 34 °C with shaking (225 rpm) for 48 hours. Samples of the cultures were then filtered and used in determination of the concentrations of sucrose, glycerol and 1,3-propanediol (PDO) in the broth by high performance liquid chromatography.

Chromatographic separation was achieved using an Aminex HPX-87P column (Bio-Rad, Hercules, CA) with an isocratic mobile phase of distilled-deionized water at a flow rate of 0.5 mL/min and a column temperature of 85 °C. Eluted compounds were quantified by refractive index detection with reference to a standard curve prepared

from commercially purchased pure compounds dissolved to known concentrations in MOPS minimal medium. Retention times were sucrose at 12.2 min, 1,3-propanediol at 17.9 min, and glycerol at 23.6 min. Table 8 shows the residual sucrose and molar yield of PDO and glycerol (mol PDO + mol glycerol/mol glucose equivalent), in the cultures of these four strains.

Table 8

Sucrose utilization and PDO and Glycerol Production

Example	Strain	44 hour sucrose g/L	Molar Yield (mol PDO + glycerol/mol glucose equivalent)
Example 18, Comparative	PDO2815 (P1.20cscB)	0.85	1.09
Example 19, Comparative	PDO2818 (P1.5cscB)	8.19	1.12
Example 20	PDO2816 (P1.20cscB17)	0.00	1.14
Example 21	PDO2819 (P1.5cscB17)	0.00	1.21

As can be seen by the results in Table 8, sucrose was completely utilized in 48 hours only in the two cultures expressing the improved facilitated diffusion transporter CscB17 (Examples 20 and 21). Furthermore, the molar yield of PDO and glycerol was greater in the cultures expressing CscB17 than in those with the wild-type sucrose symporter CscB (Comparative Examples 18 and 19). Thus, sucrose transport by facilitated diffusion was shown to be advantageous for PDO and glycerol production.

EXAMPLES 22-24

Recombinant *E. coli* Strains Comprising Variants of Sucrose Transporter Gene *scrT1* from *Citrobacter sp.* 30_2

The purpose of these Examples was to construct recombinant *E. coli* strains containing mutant transporter genes from *Citrobacter* sp. 30_2 and to demonstrate sucrose transport by facilitated diffusion. The protein encoded by the mutant sucrose transport gene was altered in a residue required for H^+ translocation, thus eliminating H^+/sucrose symport.

Construction of Expression Vectors:

Plasmid pDMWP12-scrT1, carrying a gene encoding a transporter protein from *Citrobacter* sp. 30_2, was constructed as follows. Vector pDMWP3 was obtained from Integrated DNA Technologies, Inc. (Coralville, IO). The pDMWP3 vector was constructed by cloning a promoter/ MCS/ double terminator region (set forth in SEQ ID NO:124), synthesized by Integrated DNA Technologies, Inc., into the pIDT-SMART vector (Integrated DNA Technologies, Inc.). Vector pDMWP4 was constructed from plasmid pBR322. A sca1 site and a kpn1 site on the 5' end of the *TetR* gene and an additional kpn1 site at the 3' end of the *TetR* gene were introduced into plasmid pBR322. Additionally, a kpn1 site was removed from the middle of the *AmpR* gene. All restriction sites were either added or removed using Stratagene's QuikChange® Site-Directed Mutagenesis kit (Stratagene, La Jolla, CA) following the manufacturer's protocols. Vector pDMWP3 was digested with EcoR1 and Kpn1 and the resulting 438 bp fragment was cloned into vector pDMWP4, which was also digested with EcoR1 and Kpn1, to give vector pDMWP12, which is also referred to herein as pBR*P1.20. The *scrT1* transporter gene from *Citrobacter* sp. 30_2 was codon optimized for expression in *E. coli*. The codon optimized sequence, set forth in SEQ ID NO:125, was synthesized by GenScript USA Inc. (Piscataway, NJ). The synthetic gene was subcloned into vector pDMWP12 at restriction sites of HindIII and XmaI to yield pDMWP12-scrT1. This subcloning was done at GenScript. The

presence of the transporter gene in pDMWP12-scrT1 was confirmed by sequence analysis.

The residue equivalent to R300 of *E. coli* CscB was found by multiple sequence alignment to be an arginine residue at position 305. R305 was mutated independently with two sets of primers to introduce an R305A mutation (SEQ ID NO:108, encoded by SEQ ID NO:107) and an R305L mutation (SEQ ID NO:110, encoded by SEQ ID NO:109). Site directed mutagenesis, using Stratagene's QuikChange® Site-Directed Mutagenesis kit was employed. Oligonucleotides ODMWP97 (SEQ ID NO:126) and ODMWP98 (SEQ ID NO:127) were used to introduce the R305A mutation, creating plasmid pDMWP112. Oligonucleotides ODMWP99 (SEQ ID NO:128) and ODMWP100 (SEQ ID NO:129) were used to introduce the R305L mutation, creating pDMWP113.

Construction of *E. coli* Strains Comprising the Variant Citrobacter sp. Sucrose Transporter:

Plasmids pDMWP112 and pDMWP113 were introduced into *E. coli* strain PDO3513 (described in Examples 1and 2). The resultant strains were named PDO2896 and PDO2897, respectively. Additionally, the vector pDMWP12 (described in Examples 1 and 2) was introduced into strain PDO3513 to yield strain PDO2576.

Growth Characterization of *E. coli* Strains:

E. coli strains PDO2576 (Example 22, Comparative), PDO2896 (Example 23), and PDO2897 (Example 24) were grown overnight in LB (Luria Bertani) medium containing 100 µg/mL ampicillin at 37 °C. The next day, these cultures were diluted 1:50 in LB (Luria Bertani) medium containing 100 µg/mL ampicillin. These cultures were grown at 37 °C with shaking at 250 rpm for 4 hours. The log-phase cultures were diluted 1:100 in the wells of a Bioscreen-C plate (instrument and plates purchased from Growth Curves USA, Piscataway NJ) with 150 µL MOPS minimal medium (Teknova, Half Moon Bay, CA) containing 2 g/L glucose or 8 g/L sucrose.

The cultures were grown at 37 °C in triplicate with continuous shaking and the optical density was monitored. The optical density of the cultures measured at 6 hours after inoculation is shown in Table 9.

Table 9

Optical Density of Cultures Growing on Glucose or Sucrose at 6 Hours

Carbon Source	Example 22, Comparative PDO2576	Example 23 PDO2896	Example 24 PDO2897
2 g/L glucose	0.506 ± 0.001	0.561 ± 0.021	0.569 ± 0.014
8 g/L sucrose	0.030 ± 0.002	0.439 ± 0.017	0.451 ± 0.003

As can be seen from the data in Table 9, all of the strains grew well with glucose as a sole carbon source indicating that the inoculum cultures were viable. In contrast, the control strain PDO2576 (Example 22, Comparative) was unable to grow with sucrose as sole carbon source, while the strains expressing the mutant *scrT1* genes encoding a sucrose transporter unable to translocate H^+ ion (Examples 23 and 24) were able to grow with sucrose as sole carbon source. Thus, net translocation of sucrose across the membrane must have occurred without translocation of a H^+ ion.

EXAMPLES 25 and 26

Recombinant *E. coli* Strains Comprising Variants of Sucrose Transporter Gene *scrT7* from *Bifidobacterium longum*

The purpose of these Examples was to construct a recombinant *E. coli* strain containing a mutant transporter gene from *Bifidobacterium longum* NCC2705 and to demonstrate sucrose transport by facilitated diffusion. The protein encoded by the mutant sucrose transport gene was altered in a residue required for H^+ translocation, thus eliminating H^+/sucrose symport.

Construction of Expression Vectors:

Plasmid pDMWP12-scrT7, carrying a gene encoding a transporter protein from *Bifidobacterium longum* NCC2705, was constructed using plasmid pDMWP12 (described in Examples 1 and 2 and Examples 22-24). The *scrT7* sucrose transporter gene from *Bifidobacterium longum* was codon optimized for expression in *E. coli*. The codon optimized sequence, set forth in SEQ ID NO:130, was synthesized by GenScript USA Inc. (Piscataway, NJ). The synthetic gene was subcloned into vector pDMWP12 at restriction sites of HindIII and XmaI. This subcloning was done at Genscript. The presence of the transporter gene in the vectors was confirmed by sequence analysis.

The residue equivalent to R300 of *E. coli* CscB was found by multiple sequence alignment to be an arginine residue at position 312. Plasmid pDMWP12-scrT7 was mutated to introduce an R312A mutation (SEQ ID NO:112, encoded by SEQ ID NO:111). Site directed mutagenesis, using Stratagene's QuikChange® Site-Directed Mutagenesis kit, was employed. Oligonucleotides ODMWP101 (SEQ ID NO: 131) and ODMWP102 (SEQ ID NO:132) were used to introduce the R312A mutation, creating plasmid pDMWP114.

Construction of *E. coli* Strains Comprising the Variant *Bifidobacterium longum* Sucrose Transporter:

Plasmid pDMWP114 was introduced into *E. coli* strain PDO3513 (described in Examples 1 and 2). The resultant strain was named PDO2898. Additionally, the vector pDMWP12 (described in Examples 1 and 2) was introduced into PDO3513 to yield strain PDO2576.

Growth Characterization of *E. coli* Strains:

E. coli strains PDO2576 (Example 25, Comparative) and PDO2898 (Example 26) were grown overnight in LB (Luria Bertani) medium containing 100 µg/mL ampicillin at 37 °C. The next day, these cultures were diluted 1:50 in LB (Luria Bertani) medium containing 100 µg/mL

ampicillin. These cultures were grown at 37 °C, with shaking at 250 rpm for 4 hours. The log-phase cultures were diluted 1:100 in the wells of a Bioscreen-C plate (instrument and plates purchased from Growth Curves USA, Piscataway NJ) with 150 μL MOPS minimal medium (Teknova, Half Moon Bay, CA) containing 2 g/L glucose or 8 g/L sucrose. The cultures were grown at 37 °C in triplicate with continuous shaking and the optical density was monitored. The optical density of the cultures measured at 6 hours after inoculation is shown in Table 10.

Table 10

Optical Density of Cultures Growing on Glucose or Sucrose at 6 Hours

Carbon Source	Example 25, Comparative PDO2576	Example 26 PDO2898
2 g/L glucose	0.506 ± 0.001	0.531 ± 0.011
8 g/L sucrose	0.030 ± 0.002	0.268 ± 0.005

As can be seen from the data in Table 10, both of the strains grew well with glucose as a sole carbon source indicating that the inoculum cultures were viable. In contrast, the control strain PDO2576 (Example 25, Comparative) was unable to grow with sucrose as sole carbon source, while the strain expressing the mutant *scrT7* gene encoding a sucrose transporter unable to translocate H^+ ion (Example 26) was able to grow with sucrose as sole carbon source. Thus, net translocation of sucrose across the membrane must have occurred without translocation of a H^+ ion.

What is claimed is:

1. A variant sucrose transporter polypeptide having an amino acid sequence that has at least 95% identity to an amino acid sequence as set forth in SEQ ID NO:26 based on a Clustal W method of alignment and having an amino acid change from arginine to alanine or arginine to leucine at position 300, and comprising:

(a) at least one additional amino acid change selected from the group consisting of:

(i) glutamine to histidine at position 353

(ii) leucine to proline at position 61;

(iii) phenylalanine to leucine at position 159;

(iv) glycine to cysteine at position 162;

(v) proline to histidine at position 169;

(vi) leucine to tryptophan at position 61;

(vii) leucine to histidine at position 61;

(viii) leucine to phenylalanine at position 61; and

(ix) leucine to tyrosine at position 61; or

(b) a length of 402 to 407 amino acids from the N-terminus; or

(c) a length of 402 to 407 amino acids from the N-terminus, and having at least one of the amino acid changes of (a).

2. A variant sucrose transporter polypeptide having an amino acid sequence that has at least 95% identity based on a Clustal W method of alignment to an amino acid sequence selected from the group consisting of SEQ ID NOs: 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, and 98, and comprising an amino acid at an equivalent position when compared with a reference amino acid sequence of SEQ ID NO:26 selected from the group consisting of:

(a) alanine at a position equivalent to position 300; and

(b) leucine at a position equivalent to position 300.

3. The variant sucrose transporter polypeptide of claim 2, wherein the variant sucrose transporter polypeptide further comprises:

(a) at least one of the following amino acids at an equivalent position when compared with the reference amino acid sequence of SEQ ID NO:26:

(i) histidine at a position equivalent to position 353;

(ii) proline at a position equivalent to position 61;

(iii) leucine at a position equivalent to position 159;

(iv) cysteine at a position equivalent to position 162;

(v) histidine at a position equivalent to position 169;

(vi) tryptophan at a position equivalent to position 61;

(vii) histidine at a position equivalent to position 61;

(viii) phenylalanine at a position equivalent to position 61;

(ix) tyrosine at a position equivalent to position 61; and/or

(b) truncation at a position equivalent to position 407, 406, 405, 404, 403, or 402 when compared with the reference amino acid sequence of SEQ ID NO:26.

4. A recombinant bacterium comprising in its genome or on at least one recombinant construct:

(a) a nucleotide sequence encoding a variant sucrose transporter polypeptide having an amino acid sequence that has at least 95% identity based on a Clustal W method of alignment to an amino acid sequence selected from the group consisting of SEQ ID NOs:26, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, and 98, and an amino acid at an equivalent position when compared with a reference amino acid sequence of SEQ ID NO:26 selected from the group consisting of:

(i) alanine at a position equivalent to position 300; and

(ii) leucine at a position equivalent to position 300; and

(b) a nucleotide sequence encoding a polypeptide having sucrose hydrolase activity;

wherein (a) and (b) are each operably linked to the same or a different promoter, wherein each of (a) and (b): (i) is on a recombinant construct; (ii) encodes a polypeptide

comprising a non-native polypeptide; (iii) comprises a foreign gene; or (iv) is any combination of (i), (ii) or (iii), and further wherein said recombinant bacterium is capable of metabolizing sucrose.

5. The recombinant bacterium of claim 4, wherein the variant sucrose transporter polypeptide further comprises:

(a) at least one of the following amino acids at an equivalent position when compared with the reference amino acid sequence of SEQ ID NO:26:

(i) histidine at a position equivalent to position 353;

(ii) proline at a position equivalent to position 61;

(iii) leucine at a position equivalent to position 159;

(iv) cysteine at a position equivalent to position 162;

(v) histidine at a position equivalent to position 169;

(vi) tryptophan at a position equivalent to position 61;

(vii) histidine at a position equivalent to position 61;

(viii) phenylalanine at a position equivalent to position 61;

(ix) tyrosine at a position equivalent to position 61; and/or

(b) truncation at a position equivalent to position 407, 406, 405, 404, 403, or 402 when compared with the reference amino acid sequence of SEQ ID NO:26.

6. The recombinant bacterium of claim 4 or claim 5, wherein the polypeptide having sucrose hydrolase activity is classified as EC 3.2.1.26 or EC 2.4.1.7.

7. The recombinant bacterium of claim 4 or claim 5 further comprising in its genome or on at least one recombinant construct, a nucleotide sequence encoding a polypeptide having fructokinase activity.

8. The recombinant bacterium of claim 7, wherein the polypeptide having fructokinase activity is classified as EC 2.7.1.4, EC 2.7.1.3, or EC 2.7.1.1.

9. The recombinant bacterium of claim 4 or claim 5, wherein said bacterium is selected from the group consisting of the genera: *Escherichia*, *Klebsiella*, *Citrobacter*, and *Aerobacter*.

10. The recombinant bacterium of claim 9, wherein said bacterium is *Escherichia coli*.

11. The recombinant bacterium of claim 4 or claim 5, wherein the recombinant bacterium produces 1,3-propanediol, glycerol, and/or 3-hydroxypropionic acid.

12. A process for making glycerol, 1,3-propanediol and/or 3-hydroxypropionic acid from sucrose comprising:

a) culturing the recombinant bacterium of claim 11 in the presence of sucrose; and

b) optionally, recovering the glycerol, 1,3-propanediol and/or 3-hydroxypropionic acid produced.